ESSAYS ON HEALTH, ENTREPRENEURSHIP, AND THE LABOR MARKET

by

Lulu Liu

B.S. in Economics, Peking University, 2011

M.A. in Economics, University of Pittsburgh, 2013

Submitted to the Graduate Faculty of
the Dietrich School of Arts and Sciences in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Lulu Liu

It was defended on

March 7th, 2017

and approved by

Daniel Berkowitz, Department of Economics, University of Pittsburgh

Thomas Rawski, Department of Economics, University of Pittsburgh

Werner Troesken, Department of Economics, University of Pittsburgh

Yuting Zhang, Department of Health Policy & Management, University of Pittsburgh

Dissertation Director: Daniel Berkowitz, Department of Economics, University of Pittsburgh

ABSTRACT

ESSAYS ON HEALTH, ENTREPRENEURSHIP, AND THE LABOR MARKET

Lulu Liu, PhD

University of Pittsburgh, 2017

My dissertation is at the intersection of health economics, development economics, and labor economics, with an emphasis on the labor market impact of health interventions and healthcare policies. The first chapter explores the effect of hypertension control on annual income using data from a randomized controlled trial conducted in rural China. Utilizing the exposure to the interventions as an instrumental variable for blood pressure levels, we found that a one standard deviation decrease in systolic blood pressure can significantly increase annual income by 4.8%. The second chapter investigates the impact of health insurance on entrepreneurs and estimates the alleviation of job lock following a 2009 nationwide healthcare reform—the Urban Resident Basic Medical Insurance (URBMI) program. Using 2000-2011 data from the China Health and Nutrition Survey and a difference-in-differences approach with propensity score weighting, we find that URBMI increases self-employment rate by 5.4% for the overall population. The third chapter analyzes the effect of a major policy change in China—the 2007 Property Rights Law that improved the access to credit. I found evidence that the 2007 Property Rights Law promotes entrepreneurship as measured by the number of owner-mangers in the economy, but has little impact on necessity entrepreneurs.

Keywords: Health Economics; Development Economics; Labor Economics; China.

TABLE OF CONTENTS

PREFACE . . . x

1.0 INTRODUCTION . . . 1

2.0 THE EFFECTS OF LOWERING BLOOD PRESSURE ON INCOME: EVIDENCE FROM A RANDOMIZED CONTROLLED TRIAL IN RURAL CHINA . . . 4

- 2.1 Introduction . . . 4
- 2.2 Hypertension and Institutional Background . . . 7
  - 2.2.1 Hypertension: Risk and Management . . . 8
  - 2.2.2 Usual Care in Rural China: The Three-Tier Healthcare System . . . 9
  - 2.2.3 Labor Market in Rural Qianjiang . . . 10
- 2.3 Interventions . . . 11
- 2.4 Data and Measurements . . . 12
  - 2.4.1 Measurements . . . 12
  - 2.4.2 Covariates . . . 14
- 2.5 Empirical Strategy . . . 15
  - 2.5.1 Instrumental Variable Approach . . . 15
  - 2.5.2 Addressing Concerns of Identification . . . 16
    - 2.5.2.1 Differences in Income . . . 16
    - 2.5.2.2 Instrument Validation . . . 17
    - 2.5.2.3 Health Measurement . . . 18
    - 2.5.2.4 Selection . . . 19
  - 2.5.3 Income Improvement Channels . . . 19

2.6 Results and Cost-Benefit Analysis 20
2.6.1 Instrumental Variable Results 20
2.6.2 Placebo Test Results 21
2.6.3 Subjective Measure SF36 as a Supplement 22
2.6.4 Income Improvement Channels 22
2.6.5 Cost-Benefit Analysis 23
2.7 Conclusion and Discussion 25
3.0 DOES NON-EMPLOYMENT BASED HEALTH INSURANCE PROMOTE ENTREPRENEURSHIP? EVIDENCE FROM A POLICY EXPERIMENT IN CHINA 46
3.1 Introduction 46
3.2 Institutional Background 48
3.3 Data 51
3.3.1 Data Source and Pattern 51
3.3.2 Covariates 53
3.4 Empirical Strategies 54
3.4.1 Difference-in-Differences Approach 54
3.4.2 Falsification Tests and Sensitivity Analysis 55
3.4.3 Heterogeneous Effects 57
3.5 Results 58
3.5.1 Comparing Characteristics between Study Groups 58
3.5.2 The Effect of URBMI on Self-Employment 58
3.5.3 Heterogeneous Effects 59
3.6 Discussion and Conclusion 59
4.0 ENTREPRENEURSHIP OR NECESSITY? CREDIT CONSTRAINT AND SELF-EMPLOYMENT IN CHINA 68
4.1 Introduction 68
4.2 Institutional Background 71
4.3 Theoretical Framework 72
4.3.0.1 Optimization for Self-Employers 72

4.3.0.2 Income for Wage-Earners 74
4.3.0.3 Optimal Occupational Choices 75
4.4 Data 77
4.4.1 Data Source 77
4.4.2 Two Different Types of Self-Employment 78
4.4.3 Covariates 80
4.5 Empirical Strategy 81
4.5.1 Impact on Self-Employment Propensity 81
4.5.2 Impact through Collateral Channel 82
4.6 Results 83
4.7 Discussion and Conclusion 84
5.0 BIBLIOGRAPHY 93
APPENDIX A. CLUSTER V.S. BOOTSTRAP 103
APPENDIX B. USING TWO INSTRUMENTS 104
APPENDIX C. DATA DISCUSSION 107
APPENDIX D. ALTERNATIVE EMPIRICAL STRATEGY 110
APPENDIX E. CHARACTERISTICS OF THE TWO TYPES OF SELF-EMPLOYMENT 116
E.1 A Multinomial Logistic Model 117
E.2 Income Profile and Selection 118

LIST OF TABLES

2.1 Characteristics Used to Randomly Assign Six Towns to Three Groups 35
2.2 Summary Statistics: Comparison of Characteristics at the Baseline 36
2.3 First Stage Results 37
2.4 Comparative Statistics after Propensity Score Weighting 38
2.5 Effect of Health on Percentage Change in Income: Not Disabled 39
2.6 Effect of Health on Percentage Change in Income: Farming 40
2.7 Placebo Tests: Disabled 41
2.8 Placebo Tests: Not Farming 42
2.9 Subjective Measure as Supplement: Including $\Delta SF36$ as a Covariate 43
2.10 Subjective Measure as Supplement: Instrumenting $\Delta SF36$ 44
2.11 Effect of Health on Labor Productivity and Time 45
3.1 Employment to Self-employment Transition Rates (%) by Insurance Status . 64
3.2 Comparative Statistics for City Residents with Rural and Urban Hukou . . . 65
3.3 Marginal Effect of URBMI on Self-Employment Rate 66
3.4 Falsification Tests 67
4.1 Summary Characteristics (1989-2011 Combined) 89
4.2 Impact on Self-Employment Propensity 91
4.3 Impact through Collateral Channel 92
A1 Robustness Checks for Standard Errors 103
B1 IV Regressions Using Two Instruments 105
B2 Placebo Tests Using Two Instruments 106
C1 Comparative Statistics for Different CHNS Waves 109

D1 Comparative Statistics for Individuals with and Without Non-Employment Insurance 114
D2 Alternative Strategy: Existence of Entrepreneurship Lock 115
E1 Cramer-Ridder Test 124
E2 Marginal Effect of Multinomial Logit Regression 125
E3 Transition Rate within Employment Status (%): 1989-2011 Combined 126
E4 Summary Statistics for Different Transition Status 127
E5 Self-Selection with Post-Program Estimator 128
E6 Differences in the Change of Income 129

LIST OF FIGURES

2.1 Location of Chongqing City and Qianjiang County 27
2.2 Sample Characteristics: Senior Hypertensive Patients 28
2.3 Changes in Blood Pressure with 95% CI 29
2.4 Distribution in the Percentage Change in Income 30
2.5 Percentage Change in Income with 95% CI 31
2.6 Kernel Distribution for Income in the Treatment and Control Group 32
2.7 Changes in Blood Pressure in the Treatment Group with 95% CI 33
2.8 Changes in SF36 Scores with 95% CI 34
3.1 Parallel Trends for Urban and Rural Hukou Residents 62
3.2 Heterogeneous Effect of URBMI 63
4.1 Selection into Self-Employment 85
4.2 Change in Credit Constraint 86
4.3 Self-Employment Rate (%): 1989-2011 87
4.4 Gaps in Average Years of Schooling: 1989-2011 88
4.5 NERI Financial Marketization Index in 2007 and 2008 90
C1 Frequency of Appearance in Five CHNS Waves 108
E1 Income Distribution (1989-2011 Combined) 122
E2 Income Distribution by Year 123

PREFACE

I want to thank my thesis committee members—Professor Daniel Berkowitz, Yuting Zhang, Thomas Rawski, and Werner Troesken—not only for their time and patience through this process, but also for their intellectual contributions to my development as an economist. Their encouragement and feedback have been absolutely invaluable.

A special thanks to my parents and Ricky Svoboda. Thank you all for your constant support through the ups and downs of my endeavor. It has been bumpy at times, but your confidence in me has enhanced my ability to get through it all and succeed in the end.

1.0 INTRODUCTION

My research is at the intersection of health economics, development economics, and labor economics, with an emphasis on the labor market impact of health interventions and healthcare policies. My dissertation mobilizes information on one of the world's largest economies—China—to grapple with complex issues that deservedly absorb the attention of many economists: health care, health insurance, and entrepreneurship.

The first chapter of my thesis, "The Effects of Lowering Blood Pressure on Income: Evidence from a Randomized Controlled Trial in Rural China," explores the effect of hypertension control on annual income. Health interventions can result in social welfare gains through improvement in both physical well-being and income. Previous literature encounters difficulties in estimating the causal impact due to the endogeneity of health and the prevalence of employment contracts that limit income flexibility. Studying hypertensive farmers in rural China who need to perform un-mechanized farm work to make a living creates an unusually tight link between health, productivity, and income that might be difficult to replicate in high-income societies. Hypertension is one of the most prevalent chronic diseases in the world, but little is known about the impact of high blood pressure on labor market outcomes. This paper exploits data from a randomized controlled trial designed to improve hypertension management in a rural county in southwest China and utilizes the exposure to the interventions as an instrumental variable for blood pressure levels. We find that a one standard deviation decrease in systolic blood pressure can significantly increase annual income by 4.8%, a sizable effect compared to the 2% annual increase in the control group. There is also evidence that the increase in income comes from rising labor productivity, rather than the elongation of working hours. Cost-benefit analysis shows that the gain from improved income is much larger than the cost of the interventions, suggesting that alleviating hypertension can result in large welfare gains.

Health insurance is also an important labor market consideration. Employer-provided health insurance can lower job mobility and deter entrepreneurship, creating "job lock." However, health insurance is usually endogenous, making it difficult to evaluate this barrier. In my second chapter, "Does Non-Employment Based Health Insurance Alleviate Job Lock? Evidence from a Policy Experiment in Urban China," my co-author Yuting Zhang and I investigate the impact of health insurance on entrepreneurs and estimate the alleviation of job lock following a 2009 nationwide healthcare reform—the Urban Resident Basic Medical Insurance (URBMI) program. URBMI offers health insurance to about 271 million urban residents without formal employment, many of whom were previously uninsured. We hypothesize that this new insurance option, which does not depend on formal employment, may promote entrepreneurship because it provides additional benefit out of the normal workplace. Using 2000-2011 data from the China Health and Nutrition Survey and a difference-in-differences approach with propensity score weighting, we find that URBMI increases self-employment rate by 5.4% for the overall population, eliminating the entrepreneurship barrier entirely. This effect is heterogeneous among different sub-populations: strong for individuals with up to 12 years of schooling and workers between 30 and 50 years old, but minimal for the remaining workers.

Besides health insurance, would-be entrepreneurs face many other obstacles. My third dissertation chapter, titled "Entrepreneurship or Necessity? Credit Constraint and Self-Employment in China," analyzes the affect of a major policy change in China—the 2007 Property Rights Law that improved the access to credit. I examine its impact on the formation of different types of self-employment—opportunity entrepreneurship and necessity entrepreneurship. Opportunity entrepreneurs are owner-managers who exploit business opportunities, hire workers, and promote overall economic activity. Necessity entrepreneurs are individuals who turn to self-employment as a last resort. I construct a structural model that contains both liquidity constraints and employee hiring choices to make predictions on the corresponding self-employment rates after a relaxation of credit constraints. I test the predictions and find preliminary evidence that the 2007 Property Rights Law promotes entrepreneurship as measured by the number of owner-mangers in the economy, but has little impact on necessity entrepreneurs.

The widely recognized links between health, productivity, and economic growth make it essential to critically analyze healthcare systems and their impact on labor markets. My research, centered on examining health interventions and health insurance through the lens of economic models and statistical theories, does just this. Studying the interventions and national policy changes in China reveals implications that are applicable to many other countries and opens the door to more insightful cross-institutional analyses.

2.0 THE EFFECTS OF LOWERING BLOOD PRESSURE ON INCOME: EVIDENCE FROM A RANDOMIZED CONTROLLED TRIAL IN RURAL CHINA

This chapter is coauthored with Wenxi Tang, Yan Zhang, Liang Zhang, and Yuting Zhang.

2.1 INTRODUCTION

Besides inflicting pain, poor health can also affect labor productivity and working hours, and therefore affect earnings. Health interventions, which improve both physical well-being and income, can result in social welfare gains. Understanding the causal relationship between health and income can better inform future health policies. There are two main challenges in estimating this causal relationship: first, because health is endogenous, any correlation found between health and income may be due to reverse causality or the effect of omitted variables. For example, healthier people can be more productive, which can lead to higher income; but productivity is often unobserved. In addition, individuals with higher income can in return have better access to health insurance and medical treatment and can avoid hazardous jobs, all of which leads to better health. One common omitted variable is the strenuousness of the work, which cause higher income and worse health but is hard to measure. Second, employment contracts often limit the flexible measures of income. Most papers focus on urban labor market, where wage workers can be sheltered from health shocks due to the prevalence of employment contracts. Thus there is little evidence in the literature on the effect of health on income directly, especially in the short run.

In this paper, we address the difficulties encountered by previous literature and study the labor market impact of health improvement among patients with high blood pressure. We exploit data from a randomized controlled trial designed to improve hypertension management in a rural county in southwest China. We utilize the exposure to the trial interventions as an instrumental variable to estimate the causal impact of lowering blood pressure on income. Our sample population consist of farmers who grow corn, tobacco, and other crops in the fields, harvest the crops, and sell them in local market. Since these farmers need to perform un-mechanized farm work to make a living, we observe an unusually tight link between health, productivity, and earnings that might be difficult to replicate in high-income societies.

High blood pressure is one of the most prevalent chronic diseases in the world and a main driver of the global burden of cardiovascular diseases. World Health Organization (2009) has named hypertension as the number one risk factor for mortality, with this disease being responsible for 13% of deaths globally. In China, for example, the prevalence of hypertension among adults almost doubled from 18% in 2002 to 34% in 2010 (Hou et al., 2016). Despite the prevalence of hypertension, little is known about high blood pressure?s impact on labor market outcomes.

In our data, six towns were randomly assigned into three groups: a control group with no interventions, group 1 with an integrated care model intervention, and group 2 with both the integrated care and financial interventions. Both interventions are designed to improve care delivery and reduce costs for patients with hypertension, and are successful at hypertension control, as measured by the systolic blood pressure. Thus, the trial provides an exogenous shock to the blood pressure level and we utilize the exposure to the interventions as instruments to infer causality between health and income.

Using the instrumental variables approach, we find that lowering blood pressure increases income: specifically, one standard deviation decrease in systolic blood pressure (9.54 mmHg) significantly increases income by 4.2% among farmers. Using placebo tests, we show that this effect is not driven by remittances or government subsidies. We examine the channels through which hypertension could affect income and find that the increase in income resulted from an improvement in labor productivity rather than from longer working hours. We also conduct a cost-benefit analysis showing that the gain from increased income is much larger than the cost of the interventions.

The exclusion restriction assumption of the instrumental variables approach require that both interventions are uncorrelated with the unobservables and only affect patients' income by affecting their health. Because two interventions were randomly assigned, the instrumental variables are uncorrelated with unobservables like patients' working hours and work strenuousness. Moreover, the interventions were targeted on healthcare providers rather than on patients, so the effect on patients' income can only come from changes in patients' health. We address remaining caveats related to the verification of the identification strategy in Section 2.5.2.

Our study contributes to the literature in several ways. First, estimating the causal effect of health on income is difficult because health is endogenous; our paper utilizes an instrumental variables approach and data from a randomized controlled trial that enables us to identify causality. The literature examining the relationship between health and labor market outcomes mostly focuses on correlation rather than causality, both at the macro level (Arora, 2001; Weil, 2007) and at the micro level (Bartel and Taubman, 1979). Among the papers that aim to estimate the causal effect of health on labor market outcomes, some use a dynamic programming model (Bound et al., 2010), some use panel data and individual fixed effect (Wagstaff, 2007; Pohl et al., 2014), and others use sudden health events (Halla and Zweimüller, 2013). They found large impact of health on labor market behavior.

Second, our paper is one of few studies that provides direct empirical evidence on the effect of health on direct measure of income. Because employment contracts limit the flexible measure of income, there is little evidence on the direct effect of health on income. The literature tends to measure income lost by an output decrease due to sickness and disability and evaluate outcomes such as employee disability days, early retirement, and unemployment. However, not all diseases can be well-characterized simply by absenteeism (Bleakley, 2010), and a direct measure of productivity—such as income—is more preferable. In our paper, we focus on rural self-employed farmers who need to perform un-mechanized farm work to make a living. Thus their income is very responsive to the changes in their physical well-being.

Third, our study is one of the first to measure the effect of health *improvement*, instead of health *impairment*, on income, in low or middle income countries. Most studies on the labor market impact of health focus on developed countries, by studying the incidences of severe health impairments, such as workplace accidents and road injuries (Dano, 2005; Halla and Zweimüller, 2013). However, study of health improvement, which has more direct policy implications, is rarely seen. In this paper, we fill the gaps in the literature by focusing on rural labor market in a developing country and looking at the effect of hypertension control on income. It is particularly important to study hypertension, because effective and relatively inexpensive anti-hypertension treatments have been around for more than 50 years so potential welfare gains from hypertension interventions can be large.

Fourth, we use both objective and subjective measures of health. Many previous papers use subjective measures such as self-assessed health status, that are often associated with errors. Wu (2003) argues that subjective health measures can be confounded, since an individual may simply feel better physically when the household is financially better off. We get around this issue by measuring the severity of hypertension with systolic blood pressure, which is objective and precise. Clinical trials have demonstrated that control of systolic hypertension decreases the risk of diseases such as heart failure and stroke (Izzo et al., 2000; Trialists' Collaboration et al., 2008). In addition, we use a subjective measure—RAND's SF36 score—to rule out the possibility that our findings are driven by other improvements in health besides blood pressure.

The rest of this paper is organized as follows: Section 2.2 introduces the institutional background and Section 2.3 provides the details of the randomized controlled trial. Section 2.4 presents data and measurements. Section 2.5 illustrates the empirical strategies. Section 2.6 discusses the corresponding results. Section 2.7 concludes.

2.2 HYPERTENSION AND INSTITUTIONAL BACKGROUND

Our intervention sites are located in Qianjiang county, a rural district in China with a population of 545,000. It is located 280 kilometers south east of Chongqing, one of China's four province-level municipalities (Figure 2.1). Qianjiang has 30 towns, each surrounded by around 10 villages. Of these towns, 80% are rural, and 91% of the individuals in our sample reported that they work in farming. The average annual per capita income for rural residents was 6,215 RMB (around 1,035 USD) in 2012. Qianjiang is representative of China's rural areas in terms of its population size, economic development level, and chronic disease prevalence (Tang, 2015).

2.2.1 Hypertension: Risk and Management

Hypertension is defined as having a systolic blood pressure above 140 or a diastolic blood pressure above 90 mmHg as calculated by the averaging of two or more properly measured readings (Chobanian et al., 2003).

High blood pressure, which affects nearly one billion people worldwide, is a prevalent public health problem and is the leading risk factor for cardiovascular disease globally (Wu et al., 2008). In recent years, it has become more of a serious problem in developing countries. In the middle and low income countries, the overall prevalence rate of hypertension among the general population is about 32.3% (Sarki et al., 2015). In China, the prevalence of hypertension among adults almost doubled from 18% in 2002 to 34% in 2010 (Hou et al., 2016). It now affects over 442 million people in China and is estimated to cause 20% of deaths in China (Feng et al., 2014).

Uncontrolled hypertension can cause serious damages to the heart and coronary arteries, and increase the likelihood of heart attack, stroke, kidney disease, disability, and shortened life expectancy (Stamler, 1998). Aggarwal and Khan (2006) shows that even a slightly elevated blood pressure can increase the risk of cardiovascular disease. Hypertensive crisis, systolic blood pressure rising above 180 mmHg, can also cause severe headache, shortness of breath, and severe anxiety. Ezzati et al. (2005) suggests that high blood pressure accounts for about 45% of global cardiovascular disease morbidity and mortality. In rural China, because of poor hypertension management, the rate of uncontrolled blood pressure is much higher than that in the West, estimated to be 63.4% in 2012 among Chinese with hypertension (Hou et al., 2016). In our sample, 53.4% individuals have had inpatient stays, 21.0% of which were hypertension-related.

Treatment of hypertension usually combines nonpharmacologic therapy and anti-hypertensive drug therapy. Nonpharmacologic therapy is also called lifestyle modification, including limiting dietary salt intake, limiting alcohol intake, increasing daily exercise, etc. Anti-hypertensive drugs can be classified into four categories: diuretics, calcium channel blockers (CCB), angiotensin-converting enzyme inhibitor (ACEI), and angiotensin receptor blockers (ARB). The anti-hypertensive drugs prescribed by the doctors in our trial came exclusively from the National Essential Medicine List (2012 edition) released by the China Food and Drug Administration, and all belong to one of the four categories.[1] Multiple studies have shown that all of the drugs on this list are effective in reducing blood pressure, and it is the degree of blood pressure reduction, rather than the choice of medication, that matters the most for the reduction of cardiovascular risk for hypertensive patients (Trialists' Collaboration et al., 2008). Trialists' Collaboration et al. (2014) also shows that the benefits of blood pressure-lowering drugs are proportional to the baseline cardiovascular risk. The average baseline systolic blood pressure level for the treatment groups is high—145.8 mmHg, hence the interventions can reduce cardiovascular risk sharply through ensuring medication adherence.

2.2.2 Usual Care in Rural China: The Three-Tier Healthcare System

Rural residents in China receive care from a village-town-county three-tier health care delivery system, corresponding to the three-tier administrative hierarchy (Zhang et al., 2017). Village level medical care is typically provided by an individual clinician, who sometimes sets up the clinic at his/her home. The clinician provides simple outpatient care, preventive care, and health care education to all villagers. The village clinic has a small supply of medications but often lacks rudimentary medications included in the National Essential Medicine List. At the town level, there is usually one infirmary that provides both outpatient and inpatient care. If town hospitals cannot treat the patient's illness, the patient can choose to go to hospitals in the county seat depending on his/her financial situation. As one goes up the three tiers, the health care professionals become more experienced and the services more

[1] The National Essential Medicine List (2012 edition) includes the following anti-hypertension drugs: Captopril, Enalapril, Valsartan, Sodium Nitroprusside, Magnesium Sulfate, Nitrendipine, Nifedipine, Amlodipine, Bisoprolol, Indapamide, Phentolamine, Compound Reserpine, Compound Hypoensive, and Prazosin.

expensive.

Most rural residents in China now have health insurance through the New Rural Cooperative Medical Scheme. Under the current system, most patients can go to county seat hospitals directly without going through town hospitals, but the inpatient reimbursement rate is much lower at the county level. For example, in Qianjiang in 2012, the inpatient policy reimbursement rate was 70% for services provided in the town hospitals and 55% for those in the county seat hospitals.

2.2.3 Labor Market in Rural Qianjiang

In the six towns in Qianjiang participating in our study, around 70% of the workforce works in agriculture.[2] Among senior residents, this percentage is even higher. In our sample, 90% reported their occupations to be farmers. They grow crops such as rice, corn, and tobacco, consume some as daily provision, and sell the rest as retailers on the local market. Thus their labor income depends entirely on the value of the agricultural products that they sell. Since they need to work in the fields and sell the products in the market, physical well-being can be positively related to production and income.

The major agriculture products in Qianjiang are grain (rice, corn, beans, etc.), tobacco, oil-bearing crops (peanut and rapeseed), and vegetables. In 2012 and 2013, the average annual production was 249,989 tons of grain, 8,810 tons of tobacco, 15,218 tons of oil-bearing crops, and 176,822 tons of vegetables. Overall, 33% of the grain, 79% of the vegetables, and around 80% of the tobacco and oil-bearing crops were sold as commodities on the market.[3]

For our participants, other forms of income in addition to labor income include pension and other government subsidies and family support such as remittance from their children working in the city. However, a large percentage of rural elderly are heavily relied on their own labor income and continuing to work until they are physically incapable to do so. Davis-Friedmann (1991) describe this retirement pattern as "ceaseless toil." Based on the 2009 China Health and Nutrition Survey and the 2008 China Health and Retirement Longitudinal

[2]This number is calculated using data from The China National Bureau of Statistics (2012).

[3]All the numbers were calculated using data from Chongqing Bureau of Statistics (2014). The commodity rate for each product in Qianjiang was not available so I report the average rate for Chongqing instead. The overall commodity rate in Qianjiang (69.5%) is not very different from the average in Chongqing (63.2%).

Study, Giles et al. (2011) found that at national average, over 65% of men and 40% of women still continue to work at the age of 70. The Chongqing Bureau of Statistics (2005) reported that 50.37% of Chongqing rural residents above 60 years old have labor income as their major source of income, 42.04% mainly depend on family support, and 4.81% mainly receive government support. The implementation of the New Rural Social Pension in 2012 can potentially encourage retirement among rural elderly, but the basic pension benefit is limited and our sample still shows high labor market participation rate.

In summary, for our sample—senior rural farmers—poor health can quickly have an impact on agricultural production, meaning income is highly dependent on and responsive to physical well-being.

2.3 INTERVENTIONS

This randomized controlled trial has been registered at the World Health Organization's International Clinical Trials Registry Platform (ChiCTR-OOR-14005563) and was funded by the China Medical Board (Grant # 11-069). Two interventions were implemented: an integrated care model started in August 2012, and a financial incentive contract model started in June 2013. Both interventions ended on Jun 30, 2014. Both models aim to improve hypertension management in rural China, with the integrated care model focusing on increasing the continuity of health care and adherence to medications and the financial incentive model focusing on decreasing costs of the operations. Similar models have been tested in other countries, for example, the Accountable Care Organizations in the United States, the Integrated Care Pioneers in the U.K., and the Gesundes Kinzigtal model in Germany.

In the study, six towns were randomly assigned to three groups: two treatment groups (Group 1 and Group 2) and one control group. Group 1 only received the integrated care model, and Group 2 received both the integrated care model and the financial incentive model. The control group retained its usual care. More details of the background and the interventions can be found in Zhang et al. (2017) and Tang (2015).

2.4 DATA AND MEASUREMENTS

Our sample includes 1077 patients: 283 in treatment group 1, 492 in treatment group 2, and 302 in the control group. Figure 2.2 graphs the age distribution and the baseline blood pressure distribution for all participants. The overall study population is hypertensive patients with an average age of 66 and average base systolic blood pressure of 144.6 mmHg.

2.4.1 Measurements

We use systolic blood pressure to measure the severity of hypertension. Although there is a distinction between systolic hypertension and diastolic hypertension, multiple studies show that systolic blood pressure identifies most of people needing treatment while there is a declining relative importance of diastolic pressure with advancing age (Kannel et al., 1971). In 2008, three hypertension experts proposed that the systolic reading to be the only blood pressure measurement used in tracking and diagnosing hypertension in population over age 50 (Williams et al., 2008). Thus, in our study, we focus on systolic blood pressure only. We obtained bimonthly blood pressure data from the national chronic condition management registry database and calculated the average blood pressure in 2012 and 2013 for each individual in our sample.

Since intervention 2 started in the middle of 2013, the effect of this intervention on 2013 income is limited. Thus, we do not distinguish treatment group 1 from treatment group 2, and refer to these two groups as the "treatment group."

The average change in blood pressure for the control group and the treatment group, along with 95% confidence intervals, are displayed in Figure 2.3. Subjects in the control group experienced almost no change in blood pressure, while subjects in the treatment group experienced an average blood pressure decrease of 6.19 mmHg. One common concern in the medical literature of hypertension measurement is that the office readings of blood pressure might be inaccurate, resulting in "white coat hypertension" or "masked hypertension". More specifically, white coat hypertension refers to scenarios where blood pressure is consistently elevated for office readings but not for out-of-office readings; masked hypertension refers to blood pressure being consistently elevated for out-of-office readings but not for office readings. This inaccuracy is usually caused by the pressure of being in a hospital and being treated by an unknown health care professional. In our sample, patients' blood pressure was measured repeatedly (16 times) by the same doctors, who lived in the same neighborhood, and whom the patients were familiar with. Moreover, the patients in our study had been diagnosed with hypertension at least six months earlier and had three or more measures of blood pressure greater than 140/90 mmHg in the Chinese national official medical record before our interventions began. Thus, white coat hypertension and masked hypertension were unlikely to occur in our trial.

Individuals were asked about their annual income and the change in annual income from last year in the endpoint survey, allowing us to construct income data for 2012 and 2013. One caveat is that this income includes government subsidy and pension, in addition to pure labor income. We also cannot entirely rule out remittances from their adult children who may have migrated to the city to seek jobs. To address this concern, we use the *changes* in the log of annual income as our outcome variable, and examine how the *changes* in health affect the *percentage change* in labor income. With respect to annual government subsidies and pensions, these were the same for all of the towns and did not change during our study period (Chongqing Civil Affairs Bureau, 2008). Thus the changes in the reported income should only be the result of changes in labor income and possibly remittances. We will further show that our results are not driven by remittance or government subsidies employing placebo tests illustrated in Section 2.5.2. The distribution of changes in income for the overall population is graphed in Figure 2.4. From 2012 to 2013, individuals' average income increased by 2.73%, with around 40% of people reporting no change and about 10% experiencing more than a 10% increase or decrease.

As illustrated in Section 2.2.3, most rural residents in Qianjiang have labor income, but the rest are supported by their family or the government. In order to estimate the effect on those with labor income, we focus on individuals still in the labor force. We do not have any direct indicators for their labor market status, so we use two indicators as proxies derived from two survey questions. One question asks "Are you engaged in farming?" If the answer is "Yes," then the individual is labeled as "Farming" and considered to be inside the labor force. Another question asks "Is your physical ability severely impacted by diseases?" If the answer is "No," then the individual is labeled as "Not Disabled" and listed as inside the labor force. The corresponding complementary sets are labeled as "Not Farming" and "Disabled." These two groups are largely overlapping, but there are still some discrepancies. For example, those who are not engaged in farming can still be responsible for selling products on the local market, while some do not do farm work even though their physical ability allows them to. Thus we report results using both indicators. There are 735 individuals (68%) who are not disabled and 566 individuals (53%) who are engaged in farming. Of individuals who are engaged in farming, 93% are not disabled, and of those not disabled, 72% are engaged in farming. These two groups will be the focus of our study in Section 2.5.

Figure 2.5 shows the comparison of income changes for individuals inside the labor force between the treatment group and the control group. Individuals' income in the control group increased by around 2%, while the income in the treatment group increased by about 4%.

2.4.2 Covariates

Covariates include basic individual characteristics, health related characteristics, and town level characteristics. Basic individual characteristics include age, gender, family structure (living alone, with spouse, with children only, with spouse and children, or with others), number of family members in the household, and education level (no education, elementary and middle school, high school and above). Health related characteristics are salt control, fat control, anti-hypertension drug usage, inpatient records, and personal medical expenditure. Town characteristics are the number of residents in the town and the distance from the town to county seat, taken from Table 2.1.

We show in Table 2.2 the comparative statistics of the baseline measurements. Most of the individual characteristics are statistically similar, but there are some that differ. Individuals in the treatment group have a higher baseline blood pressure and a lower baseline income than the control group. The kernel distributions of the baseline income comparing the control group and the treatment group are shown in Figure 2.6, suggesting that the income in both groups follows a regular log-normal distribution. To address the differences at the baseline, we include all the baseline values in the empirical analysis and utilize a propensity score matching (PS weighting) method.

2.5 EMPIRICAL STRATEGY

2.5.1 Instrumental Variable Approach

The most direct approach to estimate the effect of health on income is to run a regression of percentage change in income, $\Delta LogInc$, on the changes in blood pressure, ΔBP, controlling for individual characteristics at the baseline:

$$\Delta LogInc_i = \alpha_1 + \alpha_2 \Delta BP_i + \alpha_3 Cov_i + \alpha_4 logInc12_i + \alpha_5 BP12_i + \epsilon_i \tag{2.1}$$

where Cov is a vector of the covariates introduced in the last section, $logIn12$ is the baseline log of income in 2012, and $BP12$ is the baseline blood pressure in 2012. $\Delta LogInc = (logInc13 - logInc12) \times 100$, representing the percentage change in annual income.

However, using OLS regression can be problematic since health is endogenous. Thus, we used an instrumental variable strategy, of which the first stage takes the following form:

$$\Delta BP_i = \beta_1 + \beta_2 IV_i + \beta_3 X_i + \upsilon_i \tag{2.2}$$

where X is a vector of variables including Cov, $logInc12$, and $BP12$. The instrumental variable IV is a binary indicator, with 1 indicating that the individuals are in the treatment group. There are 6 towns and 55 villages in our sample. Even though theoretically standard errors should be clustered at the town level, the number of clusters are too small (Donald and Lang, 2007). We tested the size of the standard errors using different techniques: clustered at the town level, clustered at the village level, bootstrapped, and the regular robust standard error. (See Appendix A for details.) The magnitude of the standard errors does not change much, so we choose the technique that gives us the most conservative option–clustering at the village level. If the relevance restriction and exclusion restriction are both satisfied, the instrumental variable strategy can correct the reverse causality and omitted variable biases.

Thus, we can estimate the marginal causal effects of lowering blood pressure on income consistently.

For easier interpretation, we standardized the blood pressure measure. After the standardization, change in blood pressure, Δbp, had a mean of 0 and a standard deviation of 1: $\Delta bp = \frac{\Delta BP - mean(\Delta BP)}{sd(\Delta BP)}$.

Table 2.3 presents the first stage results with F-statistics for individuals inside the labor force under different specifications. The results show that the interventions significantly decreased blood pressure by 0.341-0.479 standard deviation for those not disabled and 0.302-0.378 standard deviation for those engaged in farming. All of the F-statistics were big enough that we reject the null hypothesis that IV is a weak instrument for hypertension level.

The exclusion restriction hinges on the assumption that the instrument is uncorrelated with the unobservable ($cov(IV, \upsilon) = 0$) and only affects patients' income through their health. Since the two interventions were randomly assigned, the instrumental variables were uncorrelated with unobservables like patients' working hours and work strenuousness. Since the interventions were imposed directly on health care providers rather than on patients, the effect on patients' income could only have come from changes in their health. Thus, the exclusion restriction is satisfied. We will address more concerns regarding the exclusion restriction and identification in Section 2.5.2.

We also used two different instruments, one indicating treated by intervention 1 and the other indicating treated by intervention 2. Since intervention 2 started in the middle of 2013, the effect of this intervention on 2013 income is limited. Thus we consider this approach as supplementary evidence and show the results in Appendix B. The results are qualitatively the same as our main specification using only one instrument.

2.5.2 Addressing Concerns of Identification

2.5.2.1 Differences in Income At the baseline, individuals in the treatment group have higher annual income than those in the control group. To address the unbalanced baseline income, we utilized a propensity score method with an inverse probability of treatment weighting to balance the characteristics of individuals in each group (Rosenbaum and Rubin, 1983; Hirano and Imbens, 2001). When calculating the propensity scores, we accounted for the individual-level characteristics mentioned in Section 2.4.2, together with baseline income and blood pressure, using a logistic regression to estimate the probability of being in the treatment group. We then assigned a weight to each observation that is the inverse of the estimated propensity score of the individual's assignment to its group. Weightings were calculated separately for each sub-population: individuals engaged in farming, individuals not engaged in farming, not disabled individuals, and disabled individuals. After weighting, all characteristics were balanced (Table 2.4). We report estimation results with and without propensity score weighting in Section 2.6.

If "regression to the mean" is a concern, i.e. those with higher baseline income might have a slower growth in the future, our estimation can be considered as a lower bound since treated subjects are the ones with higher income.

2.5.2.2 Instrument Validation Since randomization occurred at the town level, there are concerns that the income differences were driven by changes in non-labor income at group level, such as remittances or government subsidies, or some unobservable characteristics related to town-level crop output.

In order to rule out the possibility that the effect came from the changes in non-labor income, we conducted falsification tests using individuals outside the labor force—"Not Farming" and "Disabled"—as placebos. If the results were driven by income from remittances or government subsidies, we would observe similar effect among individuals both inside and outside the labor force; if not, the effect would only show up for individuals in the labor force.

Figure 2.7 reports the mean changes in blood pressure with 95% confidence interval for treated individuals inside and outside the labor force respectively. Blood pressure was lowered by a similar amount, regardless of their labor force status. This suggests that if their income changes are different, it is not due to the different drops in blood pressure level.

As for the unobservable characteristics related to crop output at town level, since our outcome variable is the *changes* in income, the identification is valid as long as there are no changes in elements that affected crop harvest differently for different towns. All six participating towns are close to each other, so there are limited differences in temperature, rainfall, sun exposure and other climate characteristics. Since all six towns are in the same county, there was no difference in technology adoption and profitability of different crops as well.

To further alleviate any concerns in this matter, we include town characteristics such as population and distance to county seat in the covariates to control for any differences among towns.

2.5.2.3 Health Measurement Another concern is that the results might not come directly from hypertension, but from some mental effect or other health improvement induced by the interventions. For example, getting more attention from the doctors might increase the happiness and productivity of the patients. We used RAND's Medical Outcome Study Short-Form 36-Item Health Survey (SF36) as a comprehensive subjective health measure to rule out these possibilities.[4]

The SF36 score relies upon patient self-reporting on 36 questions and it is widely used by heathcare organizations for routine monitoring and assessment of care outcomes in adult patients. For easier interpretation, we scale all the scores to be between 0 and 100. The larger the number, the healthier the individual feels. We used the total summary score (SF36 score), and the results were qualitatively the same when using more detailed indicators.

The SF36 score for each individual was obtained from the two surveys. The mean values of changes in SF36 score for individuals engaged in farming and not disabled individuals are displayed in Figure 2.8 with their 95% confidence intervals. Subjects in the control group show almost no change in SF36 score, while on average there is a 0.23 standard deviation increase for the treatment group.

We used two strategies to verify that the improvement in income was caused by decreases in blood pressure, instead of changes to other health characteristics. First, we include ΔSF—the change in SF36 score between 2012 and 2013—in the covariates and re-run the model illustrated by equations (2.1) and (2.2). If the improvement in income is indeed due to hypertension control, including ΔSF would render no difference to our main results. Second,

[4] For more details, see RAND website on SF36 Survey.

we use another instrumental variable model as below which instruments the changes in the SF36 score, ΔSF, instead of the changes in blood pressure, ΔBP:

$$\Delta sf_i = \eta_1 + \eta_2 IV_i + \eta_3 Cov_i + \eta_4 logInc12_i + \eta_5 SF12_i + \epsilon_i \tag{2.3}$$

$$\Delta LogInc_i = \mu_1 + \mu_2 \Delta sf_i + \mu_3 Cov_i + \mu_4 logInc12_i + \mu_5 SF12_i + \upsilon_i \tag{2.4}$$

where Δsf is the standardized changes in SF36 score: $\Delta sf = \frac{\Delta SF - mean(\Delta SF)}{sd(\Delta SF)}$

2.5.2.4 Selection Since the interventions are directly imposed on the delivery system and health care providers and not patients themselves, there is a potential patient selection problem. However, even though residents in these six towns have the freedom to choose their providers, they almost always go to the clinics or infirmary in their own towns because of convenience. We checked the inpatient records between 2008 and 2014, finding that only 0.8% (44 out of 5,116 cases) of town level hospital stays involved patients attending hospitals located in a different town. Thus the hospital selection problem and unintended treatment are trivial.

2.5.3 Income Improvement Channels

To further verify the second stage of the instrumental variable approach, we clarify the channels through which hypertension affects income. Since we only examine income changes over a short period of time (two years), farmers cannot improve their earnings by switching to crops that are more profitable. Thus there are only two ways for the farmers to actively increase their income: improve labor productivity or elongate working hours. Better health can potentially lead to both.

We used several indexes to measure the influence of physical well-being on different aspects of work. We have an overall index that is derived from the question "Did your physical pain negatively affect your work or housework in the past month?" It takes on values between 1 and 5, with 1 corresponding to a big impact and 5 no impact. We created three other binary indexes, one for labor time and two for labor productivity. Time index is derived from the question "Did you decrease labor hours due to health reasons in the past month?" Productivity index 1 is derived from the question "Did you feel that you can only complete part of the daily task you planned due to health reasons in the past month?" Productivity index 2 is derived from the question "Did you feel it is more difficult to work due to health reasons in the past month?" We then calculated the changes in these indexes before and after the interventions and used them as the outcome variables. Changes in the overall index ranged between -4 and 4, with a larger number indicating less of a negative impact on work due to physical pain. Changes in the time index and productivity index 1 and 2 all took the value of -1, 0, and 1. 1 indicates improved labor productivity or longer labor time compared to the previous year; -1 indicates reduced productivity or decreased time; 0 indicates no changes.

We ran the following regressions to examine whether productivity and time were affected by blood pressure level:

$$\Delta Index_i = \lambda_1 + \lambda_2 \Delta bp_i + \lambda_3 Cov_i + \lambda_4 Index12_i + \varsigma_i \tag{2.5}$$

where $\Delta Index$ is the changes between baseline and the endpoint survey in one of the four indexes, $Index12$ is the baseline index level, and Cov is the vector of covariates introduced in Section ??. If λ_4 is negative, it suggests that lower blood pressure is accompanied by smaller impact from physical pain, higher productivity, or longer working time. We conducted this analysis respectively for individuals engaged in farming, individuals who were not disabled, and the overall population.

2.6 RESULTS AND COST-BENEFIT ANALYSIS

2.6.1 Instrumental Variable Results

Table 2.5 reports the effect of changes in health on percentage changes in income for those who are not disabled. We explored different specifications and showed results both with or without PS weighting. Column (1) in the table presents estimations from an OLS regression and columns (2)-(5) the results from instrumental variable approaches. In the OLS regression, changes in systolic blood pressure level show a very small correlation with income: a one standard deviation decrease in blood pressure is related to about 1% increase in income, the equivalent of 68 RMB. After instrumenting blood pressure level, the magnitude of the effects greatly increases. For not disabled individuals, a one standard deviation decrease in blood pressure now increases income by 5.21%-5.51% without propensity score weighting (upper panel column (2)-column (5)). After balancing the baseline characteristics with propensity score weighting, the magnitude of the coefficients slightly decreases but is still very sizable: a one standard deviation decrease in blood pressure increases income by 4.20%-5.06%(lower panel column (2)-column (5)). This effect is robust under different specifications, indicating that town level differences at the baseline are not driving the results. The increase in the magnitude of the coefficient is most likely due to the omitted variable in the OLS regression, such as work strenuousness that increases income and deteriorates health at the same time.

Table 2.6 shows similar patterns for those who are engaged in farming. Simple correlation between health changes and percentage changes in income is small (1%) but the causal effect is large: a one standard deviation decrease in blood pressure increases income by 9.26%-10.29% without propensity score weighting and 7.17%-8.91% with propensity score weighting. Comparing Table 2.6 with Table 2.5, the effect is much larger for those engaged in farming than those who are not disabled. Since physical ability is most important for those engaged in farming, it is not surprising that an improvement in health benefits them the most.

The instrumental variable results reported here were calculated using a Two-Stage-Least-Square (2SLS) estimator. We also used a Limited Information Maximum Likelihood (LIML) estimator instead of the 2SLS estimator, following Imbens (2014), and the results are virtually the same.

2.6.2 Placebo Test Results

To verify our identification strategy, we conduct placebo tests using individuals outside the labor force and report. Results are reported in Table 2.7 and Table 2.8. For those whose physical abilities are severely impacted by disease (Table 2.7) or individuals not engaged in farming (Table 2.8), the coefficient estimations are all positive and insignificant, suggesting that their income stayed unchanged even though their blood pressure level dropped at a rate similar to those inside the labor force. Thus, hypertension control only improves labor income, relieving our concern that government subsidies and remittances might have confounded the results.

2.6.3 Subjective Measure SF36 as a Supplement

To rule out the possibility that the changes in income are driven by other changes brought by the interventions, instead of hypertension, we use the SF36 score as a self-assessed measure for overall health. We first include it in the covariates and report the results in Table 2.9. Including the SF36 score renders almost no change to our main results in Table 2.5 and Table 2.6. With PS weighting, a one standard deviation decrease in blood pressure at least increases income for not disabled individuals by 4.286% (upper panel column (4)) and individuals engaged in farming by 7.042% (lower panel column (4)), which is very similar to the 4.201% and 7.172% without the inclusion of changes in the SF36 score. The coefficients for the changes in SF36 score are insignificant, suggesting that all of the effects are represented by blood pressure level.

We further instrument the changes in SF36 score directly using the same instrumental variable, IV_i, which indicates subjects being treated by the interventions. Results are reported in Table 2.10, none of which are significant, suggesting that the SF36 score has no impact on changes in income. This finding supports Wu (2003)'s claim that subjective health measure can be confounded, since an individual may feel better physically and mentally when the household is financially better off.

2.6.4 Income Improvement Channels

At last, we verified the second stage of the instrumental variable regressions by examining whether health affects income through labor productivity or time. Results are reported in Table 2.11. Overall, with the decrease in blood pressure level, individuals in our study reported less negative impact on their work due to physical pain. When we examine the effect on labor productivity indexes and time index, we find that a lower blood pressure is significantly related with an increased labor productivity, but the relationship between health and labor time is not significant. These links are robust among individuals who are engaged in farming, individuals who are not disabled, and the overall population. Consistent with Tables 2.5 and 2.6, those engaged in farming are the most affected since physical ability is most important for them. This suggests that improving hypertension mainly improve earnings through increased labor productivity instead of elongated working hours.

2.6.5 Cost-Benefit Analysis

We conducted a cost-benefit analysis to assess whether these health interventions can result in welfare gains. In particular, we calculated and compared the cost of conducting the trial with the income improvement induced by the interventions.

There are four costs related to the interventions: labor costs, operational expenditures, the financial bonus, and gifts provided to survey participants. First, labor costs for extra work from health care providers in the study is an estimated 644.36 extra hours per month for all of the participating towns. The health care providers were not actually paid for these extra hours in the trial, so they have no financial incentive to exaggerate a reported number. We assigned a 21 RMB hourly wage rate in this exercise to calculate the corresponding labor cost.[5] The total cost associated with extra labor hours is thus 162,378 RMB a year ($644.36 \times 21 \times 12$ RMB). Second, the main operational expenditures were around 19,004 RMB per year. This includes costs related to staff training and group learning; printing documents for health education, health records and information transfer; and accommodation and transportation for group meetings and on-site visits. Third, the total bonus paid to the providers in Group 2 was 140,000 RMB. Even though intervention 2 started in June 2013 hence its effect on income in 2013 should be limited, we still included this cost in order to be conservative for the welfare gain estimation. Fourth, in order to increase the response rate of the in-house survey, we asked the village clinicians who know patients well to accompany the interviewers

[5] We use the average annual salary in Chongqing in 2012 as the standard to calculate the hourly wage, and assume that people work 52 weeks a year, 5 days a week and 8 hours a day. The result is 21.82 RMB/hour. Since rural area usually has a lower hourly wage than the equivalent number calculated from annual salary, this gives us an upper bound for the actual hourly wage. Also note that the minimum hourly wage in Qianjiang in 2012 is 10.5 RMB/hour (Chongqing Human Resource Bureau, 2012), so our number is twice the value.

to patient homes, and we provided some gifts such as toothpaste and soap. The value of the gifts was about 16,000 RMB. In sum, the total cost of these two interventions was 337,383 RMB per year.

When calculating the gain from the interventions, we pick the most conservative results so that our estimations are the lower-bound. Table 2.5 suggests that for not disabled individuals, one standard deviation decrease in blood pressure (9.50 mmHg) at least increases income by 4.20%, an equivalent of 287.94 RMB. Table 2.6 suggests that for those engaged in farming, one standard deviation decrease in blood pressure (9.54 mmHg) at least increases income by 7.17%, an equivalent of 490.42 RMB. On average, the blood pressure for individuals in the treatment group was lowered by 6.58 mmHg for those not disabled and 6.44 mmHg for those engaged in farming. Thus the per capita annual increase in income induced by blood pressure improvement is 199.44 RMB ($287.94 \times \frac{6.58}{9.50}$) for those not disabled or 331.06 RMB ($490.42 \times \frac{6.44}{9.54}$) for those engaged in farming. There were 4,193 hypertensive patients treated in our trial in the 4 towns. Assuming that the percentage of individuals inside the labor force in the randomized samples are the representative of all the treated 4,193 hypertensive patients, then individuals who are not disabled and individuals engaged in farming consist of 68% and 53% of the treated patients, respectively. The improved income is then estimated to be 579,698 RMB ($199.44 \times 4193 \times 68\%$) for those not disabled, or 729,511 RMB ($331.06 \times 4193 \times 53\%$) for those engaged in farming. The most conservative estimation of benefit from improved income is almost twice the size of the cost. If we take into consideration the benefit from improved physical well-being, the welfare gains could be even more substantial.

In summary, the health interventions that we conducted to lower blood pressure resulted in sizable welfare gain, even in the short run. In the long term, we expect this gain to be larger.

2.7 CONCLUSION AND DISCUSSION

In this paper, we examined the effect of hypertension control on income in rural China by taking advantage of a dataset collected from a randomized controlled trial designed to improve hypertension management. Our study participants are mainly self-employed farmers, whose physical conditions affect their labor income substantially, which allows us to observe the impact of decreasing blood pressure in the short run.

We found that a one standard deviation decrease in systolic blood pressure can significantly increase annual income by 7.17% for individuals engaged in farming and 4.20% for individuals who are not disabled. We examined the channels through which health can affect income and found that the increase in income came from an improvement in labor productivity rather than an elongation of working hours. Cost-benefit analysis shows that the gain from improved income is much larger than the cost of the interventions, suggesting that our interventions to improve hypertension resulted in substantial welfare gains.

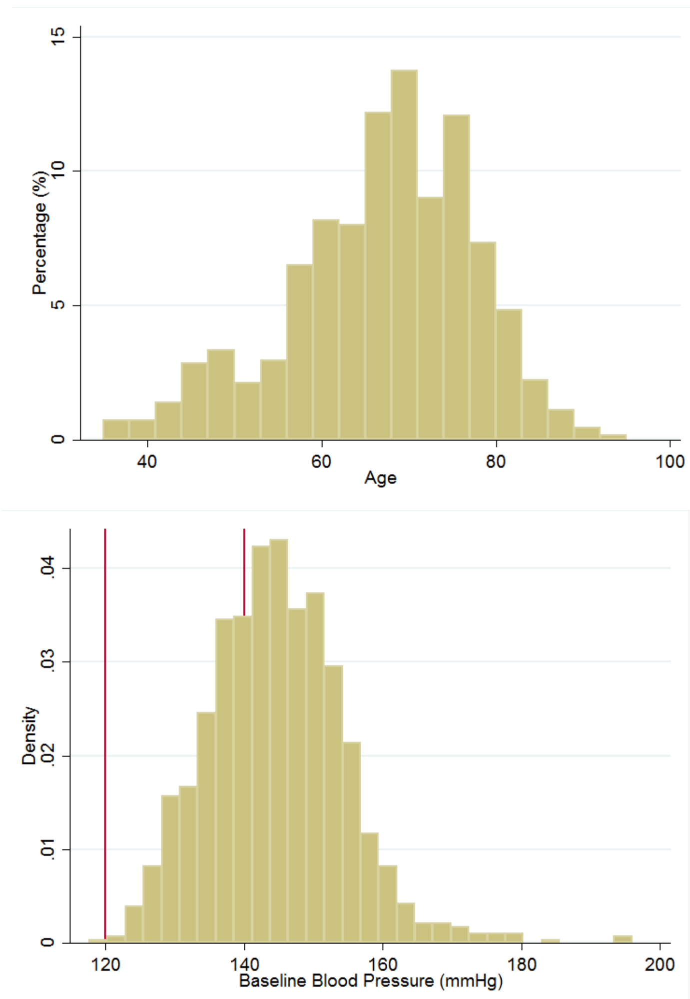
There have been rising interests in the literature on the relationship between labor market outcomes and chronic diseases, such as diabetes and depression. In general, there is a lack of research on labor market impact of all prevalent chronic diseases in low or middle income countries. To our knowledge, this paper is the first to examine the link between hypertension and income. With respect to type 2 diabetes, a literature review shows that studies on high-income countries generally find a considerable negative impact of diabetes on employment choices and income, but the economic burden associated with diabetes in developing countries is less clear (Seuring et al., 2015). In the only paper that focuses on China in this area, Liu and Zhu (2014) find that diabetes leads to an average 16.3% decrease in income after people are diagnosed. Depression is another chronic disease that has gained much attention in the economics literature. Kessler et al. (1999) suggest that treating workers' depression can help employers to save money on disability. Both Chatterji et al. (2011) and Peng et al. (2015) find that depression is associated with reductions in labor force participation and employment, but has no effect on hourly wage, most likely due to the prevalence of employment contracts (Chatterji et al., 2011; Peng et al., 2015).

Our results are encouraging and they fill the gaps in the literature on the labor market effects of chronic diseases in the developing countries, but it is not without limitations. One major concern is that the randomization was at the town level, so the results might not have been caused by lowered blood pressure but some other characteristics that varied at the town level. We alleviate this concern by noting the similarities among the six towns in agriculture technology, crop composition, and climate, and conducting placebo tests using individuals outside the labor force. Another limitation is that the findings might not be able to be extrapolated to a richer, younger, or more urban population. Our interventions were conducted in a rural county in China, where labor costs were low and hypertension management was poor. The study subjects are hypertensive farmers with an average age of 65 years old, with very different income structures compared to the urban residents. The hypertension management is more advanced in urban areas, which renders lower marginal return, while the expense of conducting similar interventions can be larger in urban areas, resulting in higher marginal cost.

However, our results can be applied to similar rural populations in other developing countries in South Asia, Africa, and South America, especially for countries where the burden of chronic conditions, including hypertension and other cardiovascular diseases, is rapidly increasing. China, the most populous country in the world and a country facing the problem of an aging population, contains 18% of the world's population. As of 2014, there are 618.7 million rural residents and 137.6 million people above 65 years old. Thus, even without extrapolating, our paper helps inform a policy discussion that is deeply important to a large segment of the world's population.

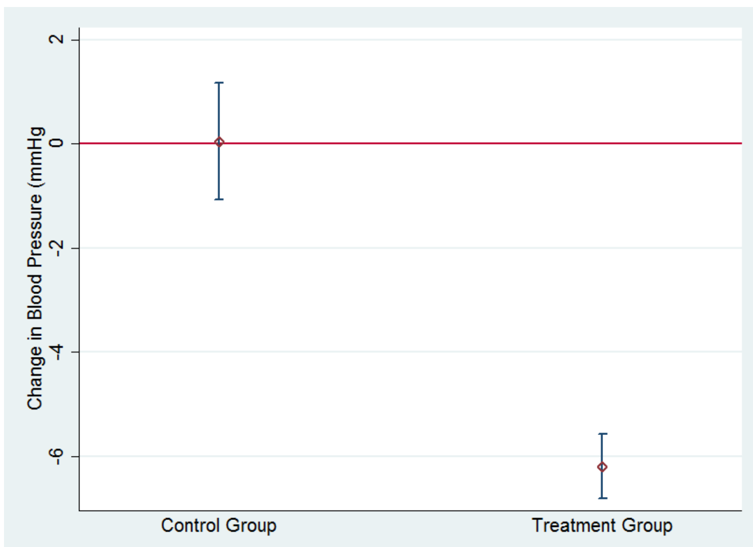
Figure 2.1: Location of Chongqing City and Qianjiang County

Figure 2.2: Sample Characteristics: Senior Hypertensive Patients



Note: Our study population is hypertensive patients with an average age of 66 and average base systolic blood pressure of 144.6 mmHg.

Figure 2.3: Changes in Blood Pressure with 95% CI



Note: Subjects in the treatment group are those exposed to intervention 1 or intervention 2.

Figure 2.4: Distribution in the Percentage Change in Income

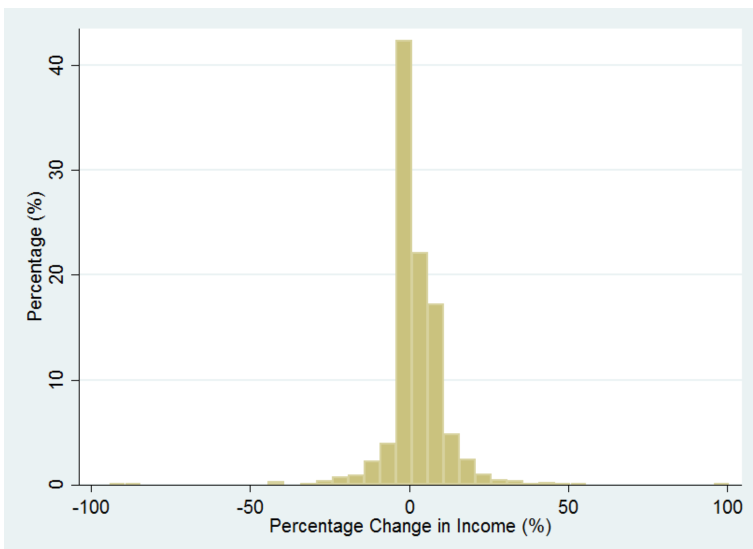
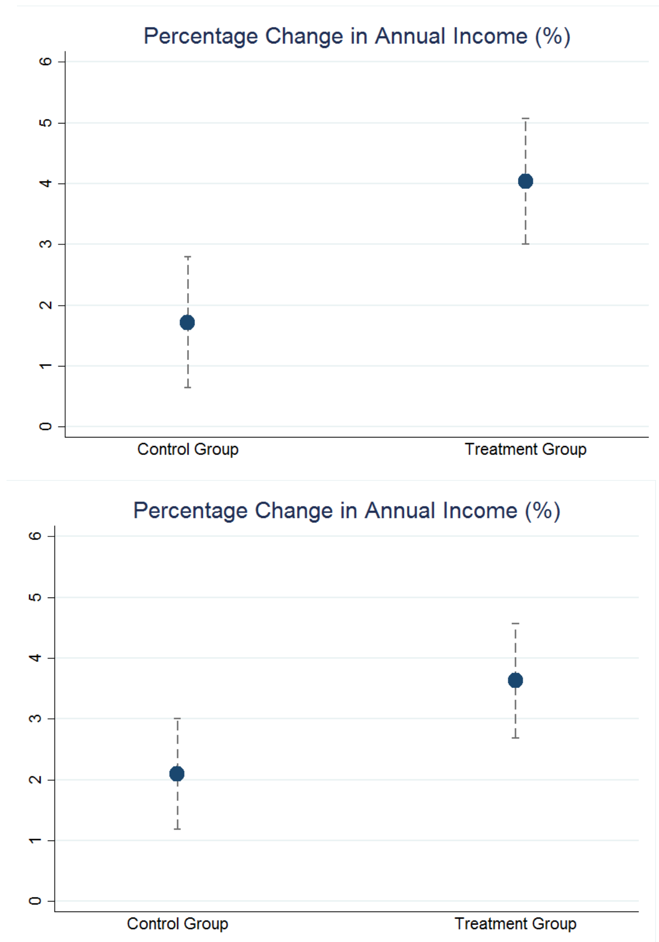


Figure 2.5: Percentage Change in Income with 95% CI



Note: Subjects in the treatment group are those exposed to intervention 1 or intervention 2. The left plot shows the percentage change in income for individulas engaged in farming and the right plot shows the number for those who are not disabled. In both cases, individuals in the treatment group have a larger growth in income.

Figure 2.6: Kernel Distribution for Income in the Treatment and Control Group

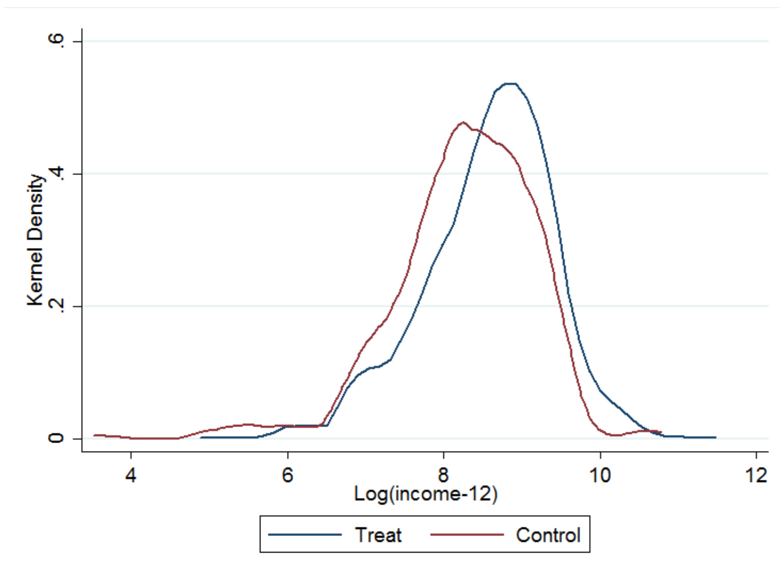
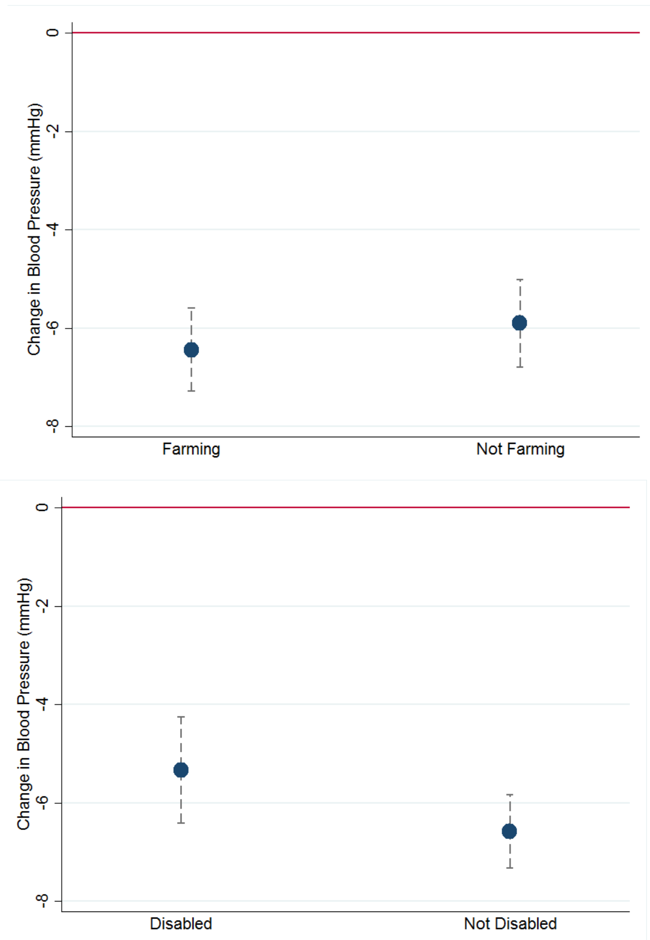
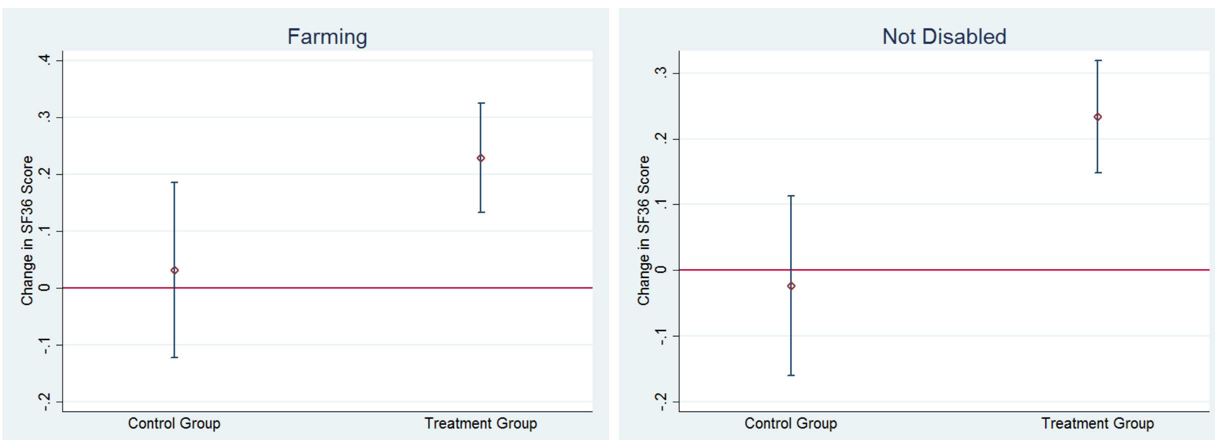


Figure 2.7: Changes in Blood Pressure in the Treatment Group with 95% CI



Note: This figure shows the changes in blood pressure level in the treatment group for those inside and outside the labor force, respectively. The left plot distinguishes individuals' labor market status by asking whether they are engaged in farming or not, and the right plot by whether they are disabled or not.

Figure 2.8: Changes in SF36 Scores with 95% CI



Note: Subjects in the treatment group are those exposed to intervention 1 or intervention 2. The left plot shows the change in SF36 score for individulas engaged in farming and the right plot shows the number for those who are not disabled.

Table 2.1: Characteristics Used to Randomly Assign Six Towns to Three Groups

	Group 1		**Group 2**		**Control**	
	Apengjiang	Jinxi	Zhoushui	Shihui	Fengjia	Shijia
Number of residents	28,000	15,600	27,055	22,448	27,464	14,126
Annual income per capita (RMB)	6452	5417	6487	5452	6900	5031
Distance to county seat (minutes)	50	90	30	60	25	100
Medical revenue (10,000 RMB)	265	169	279	198	311	159

Notes: These basic characteristics were surveyed in July 2012 and used to randomly assign towns to the three groups. We first divided the six towns equally into two clusters: one cluster of three towns with a smaller population, lower socioeconomic development levels, poorer hospital quality, and further distance away from the county seat (Jinxi, Shihui, and Shijia); and the other cluster with contrasting characteristics. In each cluster, we randomly assigned three towns to Groups 1, 2 and the control group, respectively. This table is the same as Table 1 in Zhang et al. (2017).

Table 2.2: Summary Statistics: Comparison of Characteristics at the Baseline

	mean(Control)	mean(Treat)	Difference	s.e.
Age	66.28	67.27	-0.99	0.67
Female ratio (%)	55.31	53.43	1.88	3.27
Living alone (%)	14.57	15.87	-1.30	2.46
Living with spouse (%)	36.09	32.90	3.19	3.21
Living with kids (%)	17.22	15.87	1.35	2.50
Living with spouse and kids (%)	30.13	33.81	-3.67	3.19
Attended elementary school (%)	46.95	52.36	-5.41	3.27
Attended high school or above (%)	17.04	17.34	-0.30	2.48
Family size	3.53	3.73	-0.20	0.16
Personal medical expenditure (RMB)	2566	2271	295	439
Salt control	3.22	3.42	-0.21*	0.10
Fat control	3.47	3.59	-0.11	0.10
Medicine usage (%)	35.10	36.16	-1.07	3.26
Inpatient rate (%)	18.01	18.84	-0.84	2.55
Average income in 2012 (RMB)	5482	6786	-1303***	371
Baseline Blood Pressure(mmHg)	140.87	145.79	-4.92***	0.63

Notes: Salt control and fat control: 1=never 2=occasional 3=sometimes 4=often 5=always. Significance level in t-test: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.3: First Stage Results

	(1)	(2)	(3)	(4)	(5)	(6)
Sample: Not Disabled						
IV	-0.341***	-0.429***	-0.443***	-0.387***	-0.465***	-0.479***
	(0.128)	(0.129)	(0.127)	(0.120)	(0.119)	(0.115)
Age	-0.00117	-0.00233	-0.000613	-0.00302	-0.00408	-0.00274
	(0.00419)	(0.00402)	(0.00402)	(0.00460)	(0.00450)	(0.00450)
Female ratio	-0.000673	-0.000871	-0.000950	-0.000575	-0.000717	-0.000764
	(0.000797)	(0.000744)	(0.000775)	(0.000903)	(0.000858)	(0.000877)
Log(income-12)	-0.0311	-0.0195	-0.0320	-0.0179	-0.00581	-0.0167
	(0.0352)	(0.0348)	(0.0345)	(0.0382)	(0.0402)	(0.0399)
Baseline BP	-0.0573***	-0.0598***	-0.0596***	-0.0550***	-0.0576***	-0.0571***
	(0.00378)	(0.00341)	(0.00337)	(0.00310)	(0.00267)	(0.00276)
First stage F-stats	26.35	31.77	35.71	55.49	59.89	59.49
Observations	735	735	708	735	735	708
Include Health Variables	No	No	Yes	No	No	Yes
Include Town Variables	No	Yes	Yes	No	Yes	Yes
PS Weighting	No	No	No	Yes	Yes	Yes
Sample: Farming						
IV	-0.302*	-0.354**	-0.359**	-0.324**	-0.365**	-0.378**
	(0.155)	(0.152)	(0.150)	(0.160)	(0.148)	(0.145)
Age	0.000137	-0.00208	-0.00172	0.000314	-0.00320	-0.00389
	(0.00484)	(0.00454)	(0.00460)	(0.00776)	(0.00727)	(0.00762)
Female ratio	-0.000497	-0.000676	-0.000997	-6.53e-05	-6.41e-05	-0.000245
	(0.000748)	(0.000727)	(0.000757)	(0.00116)	(0.00111)	(0.00112)
Log(income-12)	-0.00679	0.0159	0.00378	-0.00338	0.0140	0.00266
	(0.0502)	(0.0480)	(0.0482)	(0.0556)	(0.0554)	(0.0493)
Baseline BP	-0.0598***	-0.0619***	-0.0624***	-0.0620***	-0.0653***	-0.0658***
	(0.00453)	(0.00428)	(0.00447)	(0.00484)	(0.00452)	(0.00469)
First stage F-stats	22.08	22.81	25.11	21.01	21.86	23.22
Observations	566	566	541	566	566	541
Include Health Variables	No	No	Yes	No	No	Yes
Include Town Variables	No	Yes	Yes	No	Yes	Yes
PS Weighting	No	No	No	Yes	Yes	Yes

Notes: This table reports the first stage results for the IV regressions. All regressions include basic individual characteristics. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. Robust standard errors are clustered at village level. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.4: Comparative Statistics after Propensity Score Weighting

	Farming			Not Disabled			Not Farming			Disabled		
	Control	Treat	p-value	Control	Treat	p-value	Control	Treat	p-value	Control	Treat	p-value
Age	63.80	63.57	0.84	63.40	64.70	0.43	66.65	69.51	0.22	70.19	70.06	0.92
Female ratio	48.87	53.17	0.44	55.04	53.40	0.77	62.26	55.38	0.29	59.67	55.16	0.54
Live alone	12.24	11.90	0.92	12.89	13.82	0.76	20.20	20.01	0.97	29.44	20.02	0.28
Live with spouse	35.56	35.64	0.99	33.39	35.85	0.61	29.66	33.24	0.52	31.73	32.27	0.94
Live with kids	12.19	12.40	0.95	19.25	13.24	0.39	23.82	19.47	0.61	16.26	21.05	0.31
Live with spouse & kids	38.64	38.58	0.99	33.20	35.64	0.63	25.09	25.56	0.93	21.35	24.67	0.56
Attended grade school	60.24	58.19	0.69	59.03	56.22	0.60	46.67	44.60	0.79	36.73	41.71	0.48
Attended high school	20.23	20.85	0.88	17.64	19.61	0.57	10.61	13.12	0.48	9.57	12.47	0.53
Family size	3.62	3.72	0.67	3.89	3.75	0.64	3.64	3.49	0.70	3.25	3.40	0.65
Medical expenditure	1689	2082	0.37	2777	2415	0.65	3342	2741	0.43	3210	3015	0.87
Medicine usage	42.09	38.23	0.49	36.33	37.15	0.88	34.06	34.71	0.92	41.00	33.05	0.33
Salt control	3.22	3.26	0.81	3.30	3.34	0.76	3.51	3.52	0.94	3.53	3.47	0.72
Fat control	3.46	3.49	0.88	3.64	3.56	0.66	3.79	3.64	0.43	3.53	3.51	0.95
Inpatient rate	17.58	18.37	0.83	17.02	19.14	0.55	17.83	20.41	0.58	18.62	19.13	0.93
Average income in 2012	6943	6465	0.68	6834	6602	0.82	7124	6954	0.92	5240	6852	0.01
Baseline Blood Pressure	144.60	144.84	0.83	148.59	145.21	0.38	149.57	144.83	0.33	145.52	143.97	0.63

Notes: This table show the comparative statistics after propensity score weighting. Weightings are calculated separately for each group: individuals engaged in farming, individuals not engaged in farming, not disabled individuals, and disabled individuals. Individuals in the treatment groups are those exposed to the interventions. Significance level in t-test: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.5: Effect of Health on Percentage Change in Income: Not Disabled

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Without PS Weighting					
Change in blood pressure	-1.011**	-5.344*	-5.513**	-5.209*	-5.424**
	(0.394)	(2.931)	(2.781)	(2.700)	(2.628)
Age	0.0288	0.00754	0.00408	0.0292	0.0264
	(0.0499)	(0.0485)	(0.0488)	(0.0498)	(0.0501)
Female ratio	0.00468	0.00192	0.00132	0.00168	0.000940
	(0.00540)	(0.00608)	(0.00632)	(0.00641)	(0.00669)
Log(income-12)	-0.508	-0.617	-0.527	-0.801	-0.726
	(0.689)	(0.729)	(0.675)	(0.776)	(0.729)
Baseline Blood Pressure	-0.117**	-0.392**	-0.402**	-0.378**	-0.391**
	(0.0543)	(0.196)	(0.186)	(0.181)	(0.175)
With PS Weighting					
Change in blood pressure	-1.096***	-4.479**	-5.064**	-4.201**	-4.812**
	(0.345)	(1.996)	(2.294)	(1.906)	(2.208)
Age	0.0291	-0.00133	-0.00475	0.0190	0.0170
	(0.0488)	(0.0474)	(0.0481)	(0.0481)	(0.0489)
Female ratio	0.00358	0.00120	0.000704	0.00145	0.000752
	(0.00551)	(0.00622)	(0.00654)	(0.00635)	(0.00675)
Log(income-12)	-0.475	-0.480	-0.384	-0.602	-0.526
	(0.724)	(0.705)	(0.662)	(0.759)	(0.723)
Baseline Blood Pressure	-0.159***	-0.344***	-0.380***	-0.327***	-0.362***
	(0.0441)	(0.123)	(0.141)	(0.115)	(0.133)
Observations	708	735	735	708	708
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes

Notes: This table reports the effect of lowering blood pressure on percentage change in annual income (%) for individuals who are not disabled. All regressions include basic individual characteristics. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. Robust standard errors are clustered at village level. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.6: Effect of Health on Percentage Change in Income: Farming

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Without PS Weighting					
Change in blood pressure	-1.093**	-9.592*	-10.12**	-9.264*	-10.29**
	(0.422)	(5.151)	(4.963)	(4.889)	(4.904)
Age	0.0170	0.0139	-0.00395	0.0202	0.00396
	(0.0594)	(0.0627)	(0.0639)	(0.0650)	(0.0686)
Female ratio	0.00573	0.00342	0.00202	-0.000420	-0.00280
	(0.00868)	(0.0116)	(0.0122)	(0.0124)	(0.0137)
Log(income-12)	-0.689	-0.716	-0.380	-1.108	-0.771
	(0.774)	(0.892)	(0.834)	(0.945)	(0.912)
Baseline Blood Pressure	-0.113*	-0.666*	-0.704**	-0.641*	-0.707**
	(0.0660)	(0.342)	(0.331)	(0.330)	(0.330)
With PS Weighting					
Change in blood pressure	-1.312***	-7.485*	-8.873**	-7.172**	-8.915**
	(0.394)	(3.823)	(4.286)	(3.642)	(4.023)
Age	0.0146	0.0126	-0.0154	0.0101	-0.0167
	(0.0537)	(0.0665)	(0.0715)	(0.0656)	(0.0764)
Female ratio	0.00359	0.00417	0.00455	0.00190	0.00124
	(0.00887)	(0.0127)	(0.0134)	(0.0128)	(0.0140)
Log(income-12)	-0.388	-0.319	-0.0879	-0.546	-0.315
	(0.753)	(0.710)	(0.705)	(0.772)	(0.782)
Baseline Blood Pressure	-0.125*	-0.520**	-0.624**	-0.498*	-0.620**
	(0.0670)	(0.264)	(0.294)	(0.262)	(0.279)
Observations	541	566	566	541	541
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes

Notes: This table reports the effect of lowering blood pressure on percentage change in annual income (%) for individuals engaged in farming. All regressions include basic individual characteristics. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. Robust standard errors are clustered at village level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 2.7: Placebo Tests: Disabled

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Without PS Weighting					
Change in blood pressure	-0.562	1.641	2.683	2.113	2.564
	(0.458)	(3.249)	(3.585)	(3.355)	(3.444)
Age	0.0527	0.0511	0.0546	0.0384	0.0431
	(0.0386)	(0.0433)	(0.0430)	(0.0404)	(0.0386)
Female ratio	0.0126	0.0147*	0.0153*	0.0129	0.0141
	(0.00962)	(0.00820)	(0.00863)	(0.00879)	(0.00929)
Log(income-12)	-0.541	-0.0402	-0.495	-0.0554	-0.440
	(0.529)	(0.555)	(0.517)	(0.577)	(0.516)
Baseline Blood Pressure	-0.158***	-0.0201	0.0348	-0.00109	0.0247
	(0.0449)	(0.196)	(0.214)	(0.199)	(0.205)
With PS Weighting					
Change in blood pressure	-0.534	1.276	2.799	1.310	2.351
	(0.548)	(2.925)	(3.480)	(2.797)	(3.333)
Age	0.0778	0.0834*	0.0770*	0.0633	0.0597
	(0.0475)	(0.0450)	(0.0443)	(0.0437)	(0.0445)
Female ratio	0.0126	0.0146	0.0147	0.0120	0.0129
	(0.0115)	(0.0107)	(0.0112)	(0.0105)	(0.0110)
Log(income-12)	-0.424	0.0864	-0.507	0.0716	-0.447
	(0.731)	(0.795)	(0.749)	(0.754)	(0.704)
Baseline Blood Pressure	-0.157***	-0.0700	0.0373	-0.0727	0.00426
	(0.0417)	(0.178)	(0.213)	(0.168)	(0.200)
Observations	331	342	342	331	331
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes

Notes: This table reports the placebo tests using individuals who are disabled. All regressions include basic individual characteristics. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. Robust standard errors are clustered at village level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 2.8: Placebo Tests: Not Farming

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Without PS Weighting					
Change in blood pressure	-0.666	1.808	2.737	1.780	2.552
	(0.407)	(3.025)	(3.530)	(2.832)	(3.287)
Age	0.0552	0.0517	0.0561	0.0481	0.0516
	(0.0410)	(0.0474)	(0.0477)	(0.0370)	(0.0377)
Female ratio	0.00875	0.00892	0.0111	0.00764	0.0102
	(0.0109)	(0.00985)	(0.0110)	(0.0101)	(0.0114)
Log(income-12)	-0.359	0.0720	-0.206	0.105	-0.161
	(0.520)	(0.756)	(0.597)	(0.811)	(0.646)
Baseline Blood Pressure	-0.146***	0.00988	0.0552	0.00267	0.0388
	(0.0491)	(0.181)	(0.204)	(0.163)	(0.183)
With PS Weighting					
Change in blood pressure	-0.404	1.830	2.848	1.855	2.762
	(0.428)	(2.384)	(3.330)	(2.147)	(3.040)
Age	0.0552	0.0498	0.0583	0.0509	0.0541
	(0.0400)	(0.0494)	(0.0497)	(0.0373)	(0.0366)
Female ratio	0.0114	0.0123	0.0146	0.0118	0.0147
	(0.0107)	(0.0103)	(0.0123)	(0.0106)	(0.0126)
Log(income-12)	0.109	0.366	0.0557	0.367	0.102
	(0.502)	(0.596)	(0.459)	(0.625)	(0.496)
Baseline Blood Pressure	-0.117***	-0.0150	0.0581	-0.0172	0.0489
	(0.0404)	(0.137)	(0.193)	(0.119)	(0.172)
Observations	498	511	511	498	498
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes

Notes: This table reports the placebo tests using individuals not engaged in farming. All regressions include basic individual characteristics. “Health variables” include salt control, fat control, anti-hypertension drug usage, and inpatient records. “Town variables” include the number of residents in the town and the distance from the town to county seat. Robust standard errors are clustered at village level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 2.9: Subjective Measure as Supplement: Including $\Delta SF36$ as a Covariate

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Sample: Not Disabled					
Change in blood pressure	-1.069***	-4.535**	-5.068**	-4.286**	-4.824**
	(0.349)	(1.915)	(2.283)	(1.820)	(2.188)
Change in SF36 score	-0.326	-0.105	-0.0619	-0.146	-0.0934
	(0.391)	(0.394)	(0.435)	(0.397)	(0.430)
Age	0.0243	-0.00266	-0.00552	0.0168	0.0155
	(0.0484)	(0.0469)	(0.0475)	(0.0472)	(0.0479)
Female ratio	0.00328	0.00106	0.000628	0.00129	0.000654
	(0.00555)	(0.00630)	(0.00657)	(0.00644)	(0.00677)
Log(income-12)	-0.487	-0.485	-0.387	-0.608	-0.529
	(0.728)	(0.706)	(0.665)	(0.761)	(0.725)
Baseline Blood Pressure	-0.155***	-0.346***	-0.380***	-0.329***	-0.362***
	(0.0450)	(0.121)	(0.142)	(0.112)	(0.133)
Observations	708	735	735	708	708
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes
Sample: Farming					
Change in blood pressure	-1.304***	-7.384**	-8.842**	-7.042**	-8.849**
	(0.402)	(3.683)	(4.223)	(3.480)	(3.918)
Change in SF36 score	-0.120	0.254	0.295	0.267	0.350
	(0.542)	(0.637)	(0.742)	(0.644)	(0.757)
Age	0.0126	0.0153	-0.0117	0.0144	-0.0105
	(0.0530)	(0.0647)	(0.0674)	(0.0626)	(0.0696)
Female ratio	0.00336	0.00471	0.00518	0.00243	0.00194
	(0.00889)	(0.0126)	(0.0134)	(0.0127)	(0.0139)
Log(income-12)	-0.395	-0.302	-0.0681	-0.530	-0.296
	(0.757)	(0.704)	(0.713)	(0.767)	(0.791)
Baseline Blood Pressure	-0.123*	-0.516**	-0.624**	-0.492*	-0.618**
	(0.0671)	(0.258)	(0.293)	(0.254)	(0.274)
Observations	541	566	566	541	541
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes

Notes: Outcome variable is the percentage change in annual income (%). All regressions include basic individual characteristics and use PS weighting. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. "Farming" refers to individuals who are engaged in agriculture work. "Not Disabled" refers to individuals whose physical ability are not severely limited due to diseases. Robust standard errors are clustered at village level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 2.10: Subjective Measure as Supplement: Instrumenting $\Delta SF36$

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Sample: Not Disabled					
Change in SF36 score	-0.392	8.312	76.33	7.269	39.41
	(0.398)	(7.193)	(373.7)	(5.913)	(98.32)
Age	0.0268	0.104	0.953	0.126	0.624
	(0.0491)	(0.0961)	(4.653)	(0.0948)	(1.517)
Female ratio	0.00403	0.0122	0.0950	0.00933	0.0424
	(0.00561)	(0.0120)	(0.449)	(0.00982)	(0.102)
Log(income-12)	-0.475	-0.0566	3.517	-0.283	1.013
	(0.732)	(0.760)	(19.10)	(0.785)	(4.138)
Baseline Blood Pressure	-0.0958***	-0.204**	-0.836	-0.187**	-0.449
	(0.0335)	(0.0966)	(3.591)	(0.0777)	(0.870)
Observations	708	735	735	708	708
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes
Sample: Farming					
Change in SF36 score	-0.202	18.84	83.37	14.66	47.36
	(0.550)	(21.16)	(320.7)	(13.03)	(94.31)
Age	0.0166	0.211	1.029	0.245	0.831
	(0.0542)	(0.252)	(3.975)	(0.216)	(1.651)
Female ratio	0.00360	0.0443	0.184	0.0312	0.0960
	(0.00859)	(0.0500)	(0.697)	(0.0313)	(0.196)
Log(income-12)	-0.412	0.981	5.522	0.335	2.239
	(0.763)	(1.853)	(22.44)	(1.284)	(5.862)
Baseline Blood Pressure	-0.0381	-0.224	-0.603	-0.154	-0.284
	(0.0628)	(0.214)	(2.148)	(0.129)	(0.540)
Observations	541	566	566	541	541
Include Health Variables	Yes	No	No	Yes	Yes
Include Town Variables	Yes	No	Yes	No	Yes

Notes: Outcome variable is the percentage change in annual income (%). All regressions include basic individual characteristics and use PS weighting. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. "Farming" refers to individuals who are engaged in agriculture work. "Not Disabled" refers to individuals whose physical ability are not severely limited due to diseases. Robust standard errors are clustered at village level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 2.11: Effect of Health on Labor Productivity and Time

	(1)	(2)	(3)	(4)
	Δ Overall	Δ Productivity1	Δ Productivity2	Δ LaborTime
Sample: Farming				
Change in BP	-0.0870***	-0.0378*	-0.0442**	-0.00182
	(0.0337)	(0.0204)	(0.0202)	(0.0202)
Observations	568	568	568	568
Sample: Not Disabled				
Change in BP	-0.0828**	-0.0301*	-0.0281	-0.00408
	(0.0323)	(0.0178)	(0.0177)	(0.0177)
Observations	735	735	735	735
Sample: All				
Change in BP	-0.0553*	-0.0299**	-0.0244*	-0.00454
	(0.0301)	(0.0144)	(0.0140)	(0.0142)
Observations	1,077	1,077	1,077	1,077

Notes: The outcome variable for column (1) is the changes in the overall index for the influence of disease on work. It takes the value between -4 and 4. The smaller the number, the worse the influence. The outcome variables for column (2)-(4) are the changes in the time index and productivity index 1 and 2, all of which take the value of -1, 0 and 1. A value of 1 indicates improved labor productivity or longer labor time compared to the previous year; -1 indicates reduced productivity or decreased time; 0 indicate no changes. Robust standard errors are reported in parentheses. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

3.0 DOES NON-EMPLOYMENT BASED HEALTH INSURANCE PROMOTE ENTREPRENEURSHIP? EVIDENCE FROM A POLICY EXPERIMENT IN CHINA

This chapter is coauthored with Yuting Zhang.

3.1 INTRODUCTION

Previous researchers have shown that employment-based health insurance affects job mobility and deters entrepreneurship (Gruber and Madrian, 2002; Wellington, 2001; Fairlie et al., 2011; Velamuri, 2012). In particular, the fear of losing health insurance coverage discourages workers from switching jobs or starting their own business, which creates "job lock" or "entrepreneurship lock."[1] In this paper, we exploit an exogenous policy experiment in urban China to evaluate whether the introduction of a non-employer provided insurance promotes entrepreneurship.

Before 2007, the majority of working-age adults in urban China obtained their health insurance through an employer-provided insurance program, and most self-employed and unemployed were uninsured. In an effort to provide universal healthcare coverage, the Chinese government launched the Urban Resident Basic Medical Insurance (URBMI) program in 2007 in pilot cities and fully established it by the end of 2009. URBMI is a voluntary non-employer provided health insurance program that primarily provides coverage for urban residents without formal employment, including the unemployed, the self-employed, young

[1] Although "job lock" is usually a more general term, we use "job lock" and "entrepreneurship lock" interchangeably in this paper. Entrepreneurship lock refers to the barrier of becoming entrepreneurs. We are using this term to be consistent with the literature.

children and students. As of 2012, URBMI covered 271.2 million urban residents, accounting for 20.1% of the Chinese population.

We take advantage of this healthcare reform to evaluate the effect of URBMI, a non-employer sponsored insurance, on labor market outcomes, in particular, entrepreneurial decisions. We hypothesize that the implementation of URBMI increases the rate of self-employment. Before the implementation of URBMI, insurance options for self-employed individuals were limited, and workers may be reluctant to quit their jobs in order to keep their employer-sponsored health insurance. With URBMI, insurance no longer acts as a barrier that prevents workers from transitioning into self-employment.

In this paper, we examine the impact of URBMI on self-employment status, using a difference-in-differences strategy and the 2000-2011 China Health and Nutrition Survey (CHNS) data. We exploit the fact that only urban residents with urban Hukou, a legal record in the Chinese household registration system indicating formal urban residential status, were eligible for URBMI during our study period. We compare the change in self-employment propensity of urban Hukou residents following the implementation of URBMI with the concurrent change among urban residents without urban Hukou who were not eligible for URBMI. We found that URBMI increased self-employment rate by 5.4% for the overall population, 4.6% for less educated individuals (12 years of schooling or less) and 7.7% for middle age workers (between 30 and 50 years old). For more educated individuals, younger workers and senior workers, this effect is minimal.

Our study has two distinctive contributions. First, we are the first to evaluate the effects of China's new Urban Resident Basic Medical Insurance (URBMI) on labor market outcomes, in particular, entrepreneurial decisions. A few studies have formally evaluated URBMI's take-up rate and the effect on healthcare utilization, but to the best of our knowledge, no study has evaluated URBMI's effect on labor market outcomes. For instance, Lin et al. (2009) conducted a national household survey of nine representative Chinese cities, and found that there is a U-shape relationship between the URBMI participation rate and income, with the poor and those with prior hospitalizations benefiting most from URBMI. Liu and Zhao (2014) found that URBMI significantly increased the utilization of formal medical services, and Wang (2014) found that URBMI increased household-level healthcare utilization. Pan et al. (2015) concluded that URBMI beneficiaries experienced better health and received better inpatient care, than the uninsured.

Second, our paper contributes to the literature on entrepreneurship lock, by taking advantage of an exogenous policy change that disconnects employment and health insurance. Previous literature on job lock mainly focuses on employer-to-employer mobility (Madrian, 1994; Holtz-Eakin et al., 1996; Kapur, 1998), and only few studies (Fairlie et al., 2011; DeCicca, 2010; Heim and Lurie, 2010) have investigated the impact on self-employment choices. Our work shows that breaking the link between health insurance and employment can increase self-employment rates and promote entrepreneurship, but this effect varies for different populations.

The rest of this paper is organized as follows: Section 3.2 introduces the institutional background on the Chinese healthcare system. Section 3.3 presents the data source and study variables. Section 3.4 illustrates the empirical strategies. Section 3.5 shows the main results, and Section 3.6 discusses our results and concludes.

3.2 INSTITUTIONAL BACKGROUND

The Chinese health insurance system consists of three main components: the Urban Employee Basic Medical Insurance (UEBMI), the Urban Residents Basic Medical Insurance (URBMI), and the New Cooperative Medical System (NCMS). In addition to these government programs, some commercial health insurance programs exist but are in their early stages and mainly target the upper class (Dong, 2009). In 2007, less than 4% of individual medical expenses were paid by commercial health insurance in China, compared to the 50% in the United States (Wang, 2009).

UEBMI, formally launched in 1998, provides health insurance to urban residents with current or previous employment from both private and public enterprises. According to the Ministry of Health (2013), 264.7 million individuals were enrolled in UEBMI as of 2012, accounting for 19.6% of the total population in China. UEBMI is financed by premium contributions from employers (6%-8% of the employee's wage) and employees (2%-3% of their wage). Retired workers' premium contributions are fully born by their former employers (Liu, 2002). This insurance scheme has two components, an integrated social pooling fund and individual medical savings accounts. The social pooling fund is used for inpatient care. It is largely pooled at the municipal level, so the benefit level does not vary by employers. The mandated reimbursement rate varies slightly from city to city. Liang and Langenbrunner (2013) estimated the nationwide effective reimbursement rate for inpatient care was 88.5% in 2011. The individual medical savings accounts are used to cover outpatient services, deductibles, and any copayments paid by enrollees. The decision to offer UEBMI to workers is made by employers rather than employees (Xu et al., 2007). Once offered, all employees typically choose to enroll because of the generous employer subsidy.

UEBMI was designed for urban residents with official urban Hukou. Recently, some cities have expanded to make UEBMI cover migrant workers with formal employment—residents who have formal employment in the urban areas but with rural Hukou. However, the 2013 enrollment data shows that only 18.7% of the rural migrant workers had urban health insurance nationwide (Jiang and Yi, 2014), and during our study period between 2000 and 2011, the proportion was smaller than 7% on average.

URBMI was initially piloted in 2007 for selected cities and expanded to almost all cities by the end of 2009 (Liu and Zhao, 2014). This is a voluntary program that primarily provides health insurance for urban Hukou residents without formal employment, including the unemployed, the self-employed, young children and school students. According to Ministry of Health (2013), as of 2012, URBMI covered 271.2 million urban residents, accounting for 20.1% of total Chinese population. Prior to the implementation of URBMI, most of the unemployed and self-employed were uninsured, while young children and school students were usually covered by special insurance programs such as the Expanded Program of Immunization Insurance for Children.

URBMI mainly offers reimbursement for inpatient expenses, but it also covers some outpatient services for costly chronic or catastrophic diseases, like radiotherapy and chemotherapy for cancer (Liu and Zhao, 2014). Although State Council (2012) noted that the target reimbursement rate for inpatient care under URBMI was 70% in 2012, Liang and Langenbrunner (2013) estimated that the actual nationwide effective reimbursement rate in 2011 was only around 42%, because many services provided in the hospital are not reimbursable. URBMI is financed by subsidies from central and local governments (about 36% of the financing cost on average), and individual contributions (Lin et al., 2009). The annual premium for pilot cities in 2010 was between 20 and 170 CNY in central and western provinces and between 40 and 250 CNY in eastern provinces, so the program is affordable to the general public (Yip et al., 2012). Local governments have autonomy in designing and implementing URBMI according to their local needs, and therefore the local subsidy amount and reimbursement rates vary by city. Unlike UEBMI, URBMI does not have individual Medical Savings Account.

NCMS was established in 2003 to replace the old village-based rural health insurance program which targets rural Hukou residents. NCMS is a voluntary scheme, but the central and local governments provide substantial contributions to make it attractive even to low-risk households. As of 2012, the participation rate of NCMS was 98.26%, covering 805.3 million rural residents or 59.6% of the total Chinese population (Ministry of Health, 2013). Like URBMI, NCMS also mainly offers reimbursement for inpatient services, but each county can set their own benefit schemes and some affluent counties choose to cover outpatient services. In general, the level of financing is low, the coverage is shallow, but the premiums are low as well. The minimum per beneficiary contribution in 2010 was between 20 and 30 CNY in western and central provinces, and between 30 and 50 CNY in eastern provinces. Average inpatient reimbursement rate in 2010 was 43.9%, and 78.8% of the counties covered general outpatient care (Yip et al., 2012). These rates have increased substantially in recent years but we report the numbers relevant to our study period.

Recently there have been efforts to combine URBMI and NCMS to create a unified health insurance system for both rural and urban Hukou residents without employer-provided insurance (State Council, 2016). During our study period, however, there were only a few pilot cities participating in the unified program, none of which are in our sample.

In sum, before the implementation of URBMI, rural Hukou residents were mainly insured through NCMS, a non-employment based insurance program, while urban Hukou residents were largely insured through UEBMI, an employer-provided insurance program. Self-employed city residents with rural Hukou could choose to enroll in NCMS in their local towns, but self-employed city residents with urban Hukou did not have a similar government-subsidized insurance program. The latter could choose to buy commercial insurance but there was only limited supply of commercial insurance and it was usually considered unaffordable by many. After URBMI was initiated, the self-employed and the unemployed urban Hukou residents gained access to affordable health insurance through URBMI.

3.3 DATA

3.3.1 Data Source and Pattern

We use five waves of the China Health and Nutrition Survey (CHNS) data: 2000, 2004, 2006, 2009, and 2011. CHNS is a nationally representative dataset that covers both urban and rural areas, collected by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention (Zhang et al., 2014). We exclude data from rural sites because our area of focus is the urban labor market, and China's rural labor markets are dramatically different from its urban labor markets. Most rural residents are self-employed farmers, who are very different from the self-employed workers in the cities. Before 2011, this survey was conducted in nine provinces including Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, and Guizhou. In the 2011 wave, Beijing, Shanghai, and Chongqing were added. We exclude these late additions from our analysis. CHNS is a panel dataset and the same individuals were interviewed in each wave. The annual attrition rate is around 30% and new individuals were added to maintain the representativeness of the sample (See Appendix C for a detailed discussion on attrition and data representativeness).

Our main outcome is the self-employment status in each wave (1=currently self-employed; 0=wage-earners). Survey subjects report their primary and secondary employment status by choosing one of the following options: self-employed with employees, self-employed with no employees, workers for another person or enterprise, temporary workers, contract workers, paid and unpaid family business workers, or other. We define the first two categories—

self-employed with and without employees—to be self-employed individuals. Workers for another person or enterprise, temporary workers, and contract workers are categorized as wage/salary workers. The remaining categories consist only 2.7% of the entire sample, and are excluded.[2] Before the implementation of URBMI in 2000, 2004 and 2006 waves, the overall self-employment rate was 35.91%; 63.21% of wage/salary workers were covered by some type of health insurance while only 13.26% of self-employed workers were insured. After the implementation of URBMI, in our 2009 and 2011 waves, 91.72% of wage earners and 89.14% of self-employed workers were insured.

We categorize insurance types into employer-provided insurance, URBMI, commercial insurance, and other insurance.[3] Our hypothesis is that the implementation of URBMI promotes self-employment through alleviating entrepreneurship lock. In order to show the direct link between employer-provided health insurance and self-employment decisions, we focus briefly on individuals who are currently wage/salary workers in each wave, and track whether they transit to self-employment in the next wave. Table 3.1 reports the transition rate by health insurance types. We define the transition rate as the proportion of wage/salary earners in one wave who became self-employed in the next wave. This is done for all consecutive waves: 2000-2004, 2004-2006, 2006-2009, and 2009-2011. For urban Hukou residents, before the implementation of URBMI, those with employer-provided health insurance are 6.33% less likely to transition to self-employment. This difference disappears after the implementation of URBMI. For rural Hukou residents, the difference in transition rates for those with and without employer-provided health insurance is evident both before and after the implementation of URBMI. This suggests that urban Hukou residents are "treated" by URBMI and rural Hukou residents are not. To test that this pattern is not driven by other factors that only affect urban Hukou residents, we also compare the transition rates between

[2]We conduct robustness checks that allow self-employment to include paid and unpaid family business workers, or consider secondary employment as well as primary employment. The results are very similar. Thus in the rest of the paper, we only report results using primary employment and excluding family business workers.

[3]Employer-provided insurance include Worker's Compensation and Unified Planning Medical Service in 2000 and 2004, UEBMI Passway Mode, Block Model, or Catastrophic Disease Insurance in 2006, UEBMI in 2009 and 2001, and Free Medical Insurance and Unified Planning Medical Service in these five years. Other insurance include Insurance for Family Member, Cooperative Insurance, NCMS, Insurance for Women and Children, Expanded Program of Immunization Insurance for Children, and other.

individuals with and without non-employer-provided health insurance—URBMI, commercial insurance, and other insurance. There is no difference between transition rates before and after the implementation of URBMI for both urban and rural Hukou holders. This evidence in the raw data suggests that URBMI alleviated entrepreneurship lock. However, since the linkage rate for this dataset is relatively low and the linked sample is relatively small, we only consider Table 3.1 as suggestive evidence.

3.3.2 Covariates

Guided by the literature on the empirical studies of entrepreneurship (Le (1999) etc.), we control for several individual characteristics, including gender, age (and age squared), education (measured by years of schooling)[4], marriage status (1 indicating married and 0 otherwise), individual annual income, whether the spouse works, health status, occupational status, and enterprise type. To measure health status, an indicator of poor health and disability is created for respondents who indicated they suffered from one of the following conditions: heart disease, diabetes, myocardial infarction, apoplexy, asthma, cancer, goiter, angular stomatitis, blindness in one or both eye(s), loss of one or both arm(s) or leg(s), or other chronic diseases. Occupational status is measured by a categorical variable: senior professionals (doctors, professors, lawyers, engineers, etc.), junior professionals (nurses, teachers, editors, photographers, etc.), administrators/executives/managers, office staff (secretaries, office helpers, etc.), skilled workers (foremen, group leaders, craftsmen, etc.), non-skilled workers (ordinary laborers, loggers, etc.), service workers (housekeepers, cooks, waiters, doorkeepers, hairdressers, salespersons, etc.), army officers and police officers, drivers, and other occupations. These categories are taken directly from the CHNS questionnaire and are mutually exclusive. Enterprise type is measured by a categorical variable as well: state-related enterprises and non-state related. State enterprises include government departments, state services, and state-owned enterprises; non-state enterprises include private or individual enterprises, foreign invested enterprises, and others.

[4]We test alternative indicators, such as the level of completed education, and the results are similar.

3.4 EMPIRICAL STRATEGIES

3.4.1 Difference-in-Differences Approach

To study the effect of URBMI on self-employment, we utilize a difference-in-differences strategy. We take advantage of the fact that only urban residents with urban Hukou are eligible for URBMI during our study period. Urban residents with rural Hukou could enroll in NCMS throughout the entire study period, but the majority of them were not eligible to enroll in URBMI.[5] This enables us to use urban residents with urban Hukou as the treatment group, and urban residents with rural Hukou as a comparison group.

The key to our identification strategy is the *eligibility* of obtaining affordable non-employer provided health insurance, rather than the *actual status* of having health insurance. This is because the real issue with job lock is the guarantee of insurance access due to the adverse selection in the health insurance market (Baker, 2015). If the wage difference between jobs with and without employer-provided health insurance equals the cost of buying an equivalent insurance elsewhere, health insurance would have little impact on labor mobility. However, it can be difficult for workers to purchase equivalent insurance in the individual market, especially for someone with preexisting conditions. In China, this problem is more severe since the commercial insurance market was still in the early stages. Rural Hukou residents can choose to enroll in NCMS in their local towns. For urban Hukou residents without formal employment, however, affordable insurance program was only made available to them since 2007 after the implementation of URBMI. Thus, even though both URBMI and NCMS are voluntary programs, our identification is not jeopardized.

We compare the self-employment rates before and after the implementation of URBMI among urban residents with and without urban Hukou. The model specification is as follows:[6]

[5]In our sample, there are 2% of rural Hukou holders report to have URBMI. Since cities have discretion over the details of the policy, we can not rule out that no rural Hukou holders are affected by the implementation of URBMI. However, if a limited amount of rural migrants were affected, we are underestimating the effect, and the potential bias should be very small. Our model also assume that individuals do not change their Hukou status. However, about 5% of our sample did change Hukou status during our study period. We tested the models with these individuals included or excluded, and the results are very similar. We present the results with them included and use their original status as the Hukou status.

[6]It is most natural to use a probit or logit regression when the outcome variable is binary. However, using DID estimators (or any interaction terms) in a non-linear regression will result in non-consistent coefficients

$$1[Self]_{it} = \beta_0 + \beta_1 1[Post]_{it} + \beta_2 1[Hukou]_{it} + \beta_3 (1[Post]_{it} \times 1[Hukou]_{it}) \\ + \beta_4 X_{it} + \gamma_t + \tau_c + \epsilon_{it} \tag{3.1}$$

where $1[Self]_{it}$ is a binary variable with 1 indicating individual i is currently being self-employed in wave t. $1[Hukou]_{it}$ represents individual i's Hukou status in wave t: 1 corresponds to urban and 0 to rural. $1[Post]_{it}$ equals 1 for waves 2009 and 2011, and 0 for waves 2000, 2004 and 2006. X_{it} is a vector of control variables discussed in the last section. The model also includes controls for year fixed effect γ_t and city fixed effect τ_c. If the implementation of URBMI increased self-employment rate, the coefficient of the interaction term (β_3) should be positive.

We show the parallel trend between urban Hukou residents and rural Hukou residents in Figure 3.1. We run a similar regression as Equation 3.1 but without the DID terms, then plot the average values of the residuals for urban and rural Hukou residents respectively. It is clear that the two groups had similar trends before the implementation of URBMI. Rural Hukou residents had a higher self-employment rate than their urban counterparts. This is consistent with the fact that rural Hukou residents had non-employer provided health insurance (NCMS) while urban Hukou residents did not. After 2009, urban Hukou residents gained access to a non-employment based insurance scheme, and the self-employment rates of the two groups converged.

3.4.2 Falsification Tests and Sensitivity Analysis

Several potential issues arise when estimating equation 3.1. The key identification assumption is that changes in self-employment rate would be the same for residents with rural or urban Hukou in the absence of URBMI. However, there might be changes in the economy during our study period that affected self-employment propensity differently for rural and

(Ai and Norton, 2003). We calculated the marginal effect of the interaction terms in a probit model for our main results using the method in Norton et al. (2004): $\frac{\Delta^2 Prob(Self=1)}{\Delta Post \Delta Hukou} = \Phi(\beta_1 + \beta_2 + \beta_3 + Cov\beta) - \Phi(\beta_1 + Cov\beta) - \Phi(\beta_2 + Cov\beta) + \Phi(Cov\beta)$ where $Cov\beta = \beta_0 + \beta_4 X + \gamma_t + \tau_c$. Similar results are obtained, which suggests that any biases related to the use of linear models are likely to be small. Thus, we only report results from linear probability model since it is much less cumbersome to report.

urban Hukou residents. In addition, rural Hukou holders may not be comparable to urban Hukou holders. We use several strategies to address these concerns.

First, we utilize a propensity score method with an inverse probability of treatment weighting (IPTW) to balance the characteristics of urban residents with rural and urban Hukou (Rosenbaum and Rubin, 1983; Hirano and Imbens, 2001; Hirano et al., 2003). In calculating the propensity scores, we account for individual-level characteristics mentioned above using a logistic regression to estimate the probability of being urban Hukou holders. We then assign a weight to each observation that is the inverse of the estimated propensity score of the individual's assignment to its group. We report estimation results with and without propensity score weighting in Section 3.5.

Second, we include individual fixed effects to control for unobserved time-invariant individual characteristics. We report estimation results with and without individual fixed effect in Section 3.5 as well.

Third, we add to our covariates provincial macroeconomic indicators that may affect self-employment behaviors for each corresponding year. These indicators include the gross regional product of the province, urban population size, unemployment rate, value-added of production and manufacture industry, and the value of import and export of the province. We only use provincial level variables rather than a finer geographic classification, since city information is censored due to privacy concerns. We combine data from China National Statistics Bureau and CHNS to conduct our analysis.[7] In this model, standard errors are clustered at the provincial level.

Forth, we are concerned that changes in self-employment propensity may be driven by changes in the composition of rural Hukou holders, since some of the rural Hukou holders are migrants from rural areas. To formally test for this potential bias, we follow methods in Rossin-Slater et al. (2013) and run several falsification tests. We estimate models similar to our main specification but replace the original outcome variable, self-employment status, with the following outcomes: age, education, logarithms of individual annual income, and

[7]There are some concerns about the quality of China's macroeconomics statistics, especially at the provincial level. Provinces have incentives to over-report their economic performances. Rawski (2001) found that the sum of provincial GDP data is larger than the national data calculated by China National Statistics Bureau. However, we include in our model year and city fixed effect, so this bias is greatly mitigated.

gender. If the interaction term does not have a significant effect on these outcome variables, it suggests that our main results are not being driven by compositional changes due to selected migration.

Finally, we include an alternative empirical strategy that only looks at urban Hukou individuals, and show that job lock existed before 2009 but disappeared after 2009 (see details in Appendix D). We focus on the transition from a salaried position to self-employment, and use two difference-in-differences models separately for the two periods before and after the implementation of URBMI. We estimate the entrepreneurship barrier using a strategy following Holtz-Eakin et al. (1996) and Fairlie et al. (2011): comparing the effect of employment-based insurance on self-employment transition rate between those with low and high demand for the employment-based health insurance. Because the sample size is significantly smaller when we only focus on individuals who transitioned from salaried positions to self-employment, we do not report this as main results and only illustrate the details of this strategy and report its results in the appendix.

3.4.3 Heterogeneous Effects

To show the heterogeneous effects of this policy change, we conduct analysis for several subgroups—males v.s. female, more educated individuals (> 12 years of education) v.s. less educated (≤ 12), and young workers (aged < 30) v.s. middle age workers (aged 30-50) v.s. senior workers (aged > 50). We hypothesize that individuals should be affected by the implementation of URBMI less if their health care costs are relatively small compared to their wage income. Thus, the estimated marginal effect (β_3) for younger workers and more educated individuals should be smaller, since younger workers have smaller health care cost and more educated individuals usually have higher wage income.

3.5 RESULTS

3.5.1 Comparing Characteristics between Study Groups

Summary statistics of covariates for urban and rural Hukou residents are shown in Panel A of Table 3.2.[8] Panel A shows that, on average, rural Hukou individuals are less educated, have lower income, and are much more likely to be non-skilled workers and service workers. We use a propensity score method with an inverse probability of treatment weighting, among other strategies, to address this issue. We compare the weighted characteristics for each group and report them in Panel B of Table 3.2. After weighting, all key characteristics are balanced at 5% significance level.

3.5.2 The Effect of URBMI on Self-Employment

Table 3.3 shows the main results from the difference-in-differences model that compares self-employment rate among urban residents with and without urban Hukou, in the post-policy period relative to the pre-policy period. The marginal effect of the implementation of URBMI on self-employment status ranges from 5.36% to 8.34% under different specifications for the overall population. This suggests that the implementation of this non-employment based insurance increases more self-employment activities for urban Hukou residents by at least 5.36 percentage points, compared to rural Hukou residents. The baseline self-employment rate was 35.9 percentage points, resulting in a 14.9% (5.36/35.9) URBMI related self-employment rate increase.

Without the inclusion of propensity score weighting or individual fixed effect, the magnitude of the marginal effect is 7.42%, which is slightly larger than the magnitude with the inclusion of propensity score weighting (5.39%) or individual fixed effect (5.36%). Comparing column (1) and (2), we can see that the magnitude is smaller with the inclusion of provincial level macroeconomic indicators. Overall, the results are robust and statistically significant.

We present the results of falsification tests in Table 3.4. We conduct these tests in order

[8] We report the population share of certain occupation for the seven major categories that contain at least 5% of the population, since it is cumbersome to report all categories.

to rule out the possible endogeneity issue due to selected migration from rural areas to the city. None of the coefficients of the interaction term $1[Post]_{it} \times 1[Hukou]_{it}$ are significant, which indicates that the main results in Table 3.3 are not driven by compositional changes due to selected migration.

3.5.3 Heterogeneous Effects

We examine the heterogeneous effects of this policy change and conduct sensitivity analysis for different sub-groups. Compared to female workers, male workers received a larger impact from the implementation of URBMI: 6.81% with propensity score weighting and 6.30% with individual fixed effect, as shown in Table 3.3. Results for other subgroups are shown in Figure 3.2. Consistent with our hypothesis, only the self-employment propensities of less educated individuals (high school and below) are affected by the implementation of URBMI (4.56%), but more educated individuals are not affected at all. Among different age groups, only middle age workers (between 30 and 50 years old) are significantly affected by this policy change (7.73%). For senior workers, the effect is not significant, and for young workers, the point estimate is negative.

3.6 DISCUSSION AND CONCLUSION

We find strong evidence that the implementation of a non-employment based health insurance, URBMI, promotes entrepreneurship in the urban labor market in China. Comparing urban residents with and without urban Hukou, we find that the implementation of URBMI increased the self-employment rate by 5.4 percentage points, or a relative 14.9% for the overall population. The effect of URBMI varies for different sub-groups. For more educated individuals, younger workers and senior workers, there are minimal changes in self-employment activities. For less educated individuals and middle age workers, URBMI significantly increased their self-employment propensities. The heterogeneous effects are consistent with our hypotheses and with previous literature.

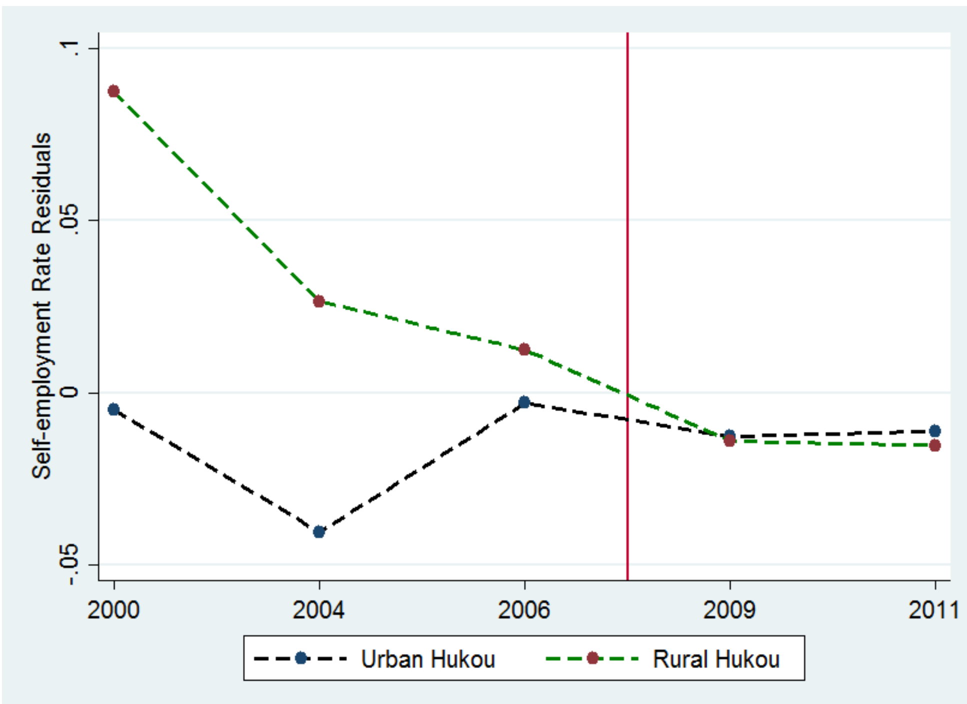
Hamersma and Kim (2009) shows that the magnitude of job lock potentially decreases with income, thus the effect of a non-employer provided health insurance program would increase self-employment propensity more for lower wage income earners. This is because employment benefits, such as employer-provided health insurance, can have larger effects on job mobility for workers with lower wage payment. Since we only have individuals' current income, but not their potential wage income, we use education level as an indicator. The result is consistent: more educated individuals, potentially with higher wage income, are not affected by the implementation of URBMI, while less educated ones are.

We also observe that only middle age workers (between 30 and 50 years old) are significantly affected by this policy change. It is not surprising that employer-provided health insurance has little impact on younger workers because their health costs are low. For older workers, although their health cost is high, a non-employer provided health insurance would mainly affect their decision to retire, rather than switching between wage sector and self-employment sector.

The U.S. government also has taken a number of measures to mitigate job lock, like Consolidated Omnibus Reconciliation Act of 1985 (COBRA) and Health Insurance Portability and Accounting Act in 1996 (HIPPA). Papers examining the effect of these policies reached different conclusions. Bailey (2013) found that the expansion of dependent coverage under the Affordable Care Act did not increase job mobility for young adults. On the other hand, Fairlie et al. (2011) found a nearly 3.1% increase in business ownership that can be explained by the new Medicare eligibility when individuals turn 65. DeCicca (2010) also suggested that the implementation of New Jersey's 1993 Individual Health Coverage Plan that partially disconnected health insurance and employment increased self-employment rate. The results seem contradicting, but they can be explained by the heterogeneous effects of job lock, consistent with what we showed in our paper. Job lock can be a minor issue for younger workers who are the study population in Bailey (2013), but it can be a major concern for less privileged population and workers who bear higher health cost as in Fairlie et al. (2011).

Our study is one of the few papers (Lin et al., 2009; Liu and Zhao, 2014; Wang, 2014; Pan et al., 2015) to evaluate the effects of the China's new Urban Resident Basic Health Insurance plan, and the first to look at its impact on entrepreneurship, or labor market outcome in general. By studying this exogenous quasi-natural policy experiment, we also contribute to the literature on job lock caused by employment-based health insurance. We conclude that breaking the link between employment and health insurance can promote entrepreneurship, an outcome that could inform policy making in many other countries.

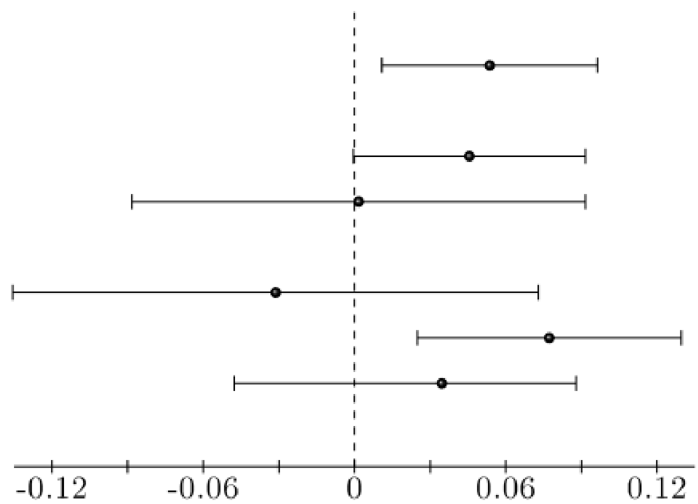
Figure 3.1: Parallel Trends for Urban and Rural Hukou Residents



Note: We run a similar regression as Equation 3.1 but without the DID terms, then plot the average values of the residuals for urban and rural Hukou residents respectively.

Figure 3.2: Heterogeneous Effect of URBMI

	No.	Estimation
All Population	7801	0.0536**
Education:		
Less Educated	6531	0.0456*
More Educated	1270	0.0017
Age:		
Young	1341	-0.0313
Middle Age	4537	0.0773**
Senior	1923	0.0201



Note: This table reports treatment effects of the interaction term on self-employment status for different sub-groups, together with the 95% confidence intervals. Model specification is the one with individual fixed effect and with provincial level macro covariates. "More educated" refers to individuals with more than 12 years of schooling. "Young" refers to individuals below 30 years old, and "Senior" refers to above 50 years old. Robust standard errors are clustered at the province level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 3.1: Employment to Self-employment Transition Rates (%) by Insurance Status

	Urban Hukou		Rural Hukou	
	Pre-policy	Post-policy	Pre-policy	Post-policy
Panel A: Employer-provided Health Insurance (EPHI)				
(1) With EPHI	2.76	4.08	3.13	0.00
(2) Without EPHI	9.09	6.25	28.13	20.00
(3) Difference: (2)-(1)	6.33***	2.17	25.00***	20.00**
Panel B: Non Employer-provided Health Insurance (Non-EPHI)				
(4) With Non-EPHI	3.53	7.14	20.59	19.20
(5) Without Non-EPHI	5.54	3.94	26.60	11.11
(6) Difference (5)-(4)	2.01	3.20	6.01	-8.09
Total Counts	1449	518	264	145

Notes: Transition rate is defined as the proportion of new self-employed individuals among the original wage-workers. Pre-policy refers to 2000-2004, 2004-2006 and 2006-2009 panels; Post-policy is the 2009-2011 panel. We conduct t-test for comparison between pre and post policy period in row (3) and (6). Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.2: Comparative Statistics for City Residents with Rural and Urban Hukou

	(1) Rural Hukou	(2) Urban Hukou	(3) Difference	(4) Observations
Panel A: Without Weighting				
Age (year)	41.60	40.83	0.78**	8268
Female Share (%)	46.20	43.43	2.77*	8280
Education (year)	7.24	11.18	-3.93***	7904
Marriage Rate (%)	85.20	82.05	3.15***	8197
Poor Health (%)	15.36	16.01	-0.65	8280
Log(Income)	7.23	8.75	-1.52***	8280
Spouse Works (%)	67.61	57.13	10.48***	8280
Senior Professional (%)	0.60	11.61	-11.00***	8265
Junior Professional (%)	1.61	8.98	-7.37***	8265
Administrator (%)	2.48	11.34	-8.86***	8265
Office Staff (%)	2.28	13.14	-10.86***	8265
Skilled Worker (%)	6.00	13.63	-7.64***	8265
Non-skilled Worker (%)	11.16	11.27	-0.11	8265
Police or Army (%)	0.23	0.97	-0.73***	8265
Driver (%)	3.72	4.07	-0.35	8265
Service Worker (%)	11.90	17.25	-5.35***	8265
Panel B: With Inverse Probability of Treatment Weighting				
Age (year)	40.38	41.28	0.90*	8268
Female Share (%)	43.57	43.79	0.22	8280
Education (year)	9.64	9.60	-0.04	7904
Marriage Rate (%)	80.10	82.88	1.78*	8197
Poor Health (%)	15.52	15.80	0.28	8280
Log(Income)	8.16	8.17	0.01	8280
Spouse Works (%)	56.27	59.98	3.71*	8280
Senior Professional (%)	5.35	7.41	2.06	8265
Junior Professional (%)	6.38	6.10	-0.72	8265
Administrator (%)	8.16	8.11	-0.05	8265
Office Staff (%)	9.61	9.05	-0.56	8265
Skilled Worker (%)	10.60	10.50	-0.10	8265
Non-skilled Worker (%)	11.63	11.30	-0.27	8265
Police or Army (%)	0.56	0.63	-0.07	8265
Driver (%)	3.99	3.93	-0.06	8265
Service Worker (%)	16.27	15.73	-0.54	8265

Notes: All samples are for adult residents living in urban areas. Panel A reports results without the weighting, and Panel B reports results with inverse probability of treatment weighting. Significance level in t-test: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 3.3: Marginal Effect of URBMI on Self-Employment Rate

	(1)	(2)	(3)	(4)
UrbanHukou*Post with Individual FE	0.0536**	0.0649***	0.0630*	0.0361*
	(0.0218)	(0.0175)	(0.0297)	(0.0189)
UrbanHukou*Post with IPTW	0.0539	0.0623**	0.0681	0.0425
	(0.0473)	(0.0270)	(0.0572)	(0.0630)
UrbanHukou*Post	0.0742**	0.0834***	0.0766*	0.0694*
	(0.0315)	(0.0150)	(0.0357)	(0.0354)
Observations	7,801	7,801	4,359	3,442
Baseline Dependent Mean	0.3591	0.3591	0.3354	0.3873
Macro Variables	Yes	No	Yes	Yes
Sample	All	All	Male	Female

Notes: Dependent variable is $SELF$, an indicator for self-employment. This table reports treatment effects of the interaction term on self-employment status, with inverse probability of treatment weights using propensity scores (IPTW), with individual fixed effect (Individual FE), and in the absence of weighting and individual fixed effect. All specifications include year and city fixed effects. "Macro Variable" include gross regional product of the province (in 100 million RMB), urban population size, unemployed population, value-added of second industry, and the value of import and export from that province. For models that include "Macro Variable", robust standard errors are clustered at the province level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table 3.4: Falsification Tests

	(1)	(2)	(3)	(4)
	Age	Education	log(Income)	Female
UrbanHukou*Post	-0.416	-0.0144	-0.316	-0.0229
	(0.596)	(0.151)	(0.237)	(0.0213)
Observations	7,801	7,801	7,801	7,801

Notes: This table presents the treatment effect of the interaction term (having urban Hukou and in post policy period) on various outcome variables as falsification tests. The estimated models are similar to our main specification in equation 3.1, but with different outcome variables. Robust standard errors are in parenthesis. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

4.0 ENTREPRENEURSHIP OR NECESSITY? CREDIT CONSTRAINT AND SELF-EMPLOYMENT IN CHINA

4.1 INTRODUCTION

According to traditional economic models, entrepreneurship spurs innovation, creates jobs, and promotes overall economic activity (Schumpeter, 1934). More recently, however, economists have come to suggest that traditional thinking about entrepreneurial activity is overly optimistic, particularly in developing countries, where often people turn to self-employment when other employment opportunities are scarce or non-existent (Acs, 2006). Self-employment rate, the most common measures of entrepreneurship, fails to distinguish different types of entrepreneurs. A self-employed worker can be an *opportunity entrepreneur*, who fits the traditional economic model, exploits new opportunities and improves product; or he/she can be a *necessity entrepreneur* or *entrepreneur by default*, who uses self-employment as a last resort when regular jobs become scarce or pay poorly. The importance of this distinction lies in its welfare implication. While a rising population of opportunity entrepreneurs may accelerate economic growth, higher rate of necessity entrepreneurs can simply indicate that the economy is creating too few conventional wage-earning jobs. Thus, it is essential to discuss the effect of some commonly used policy tools on different types of entrepreneurs, such as improving the access to credit and financial market.

Credit constraint is often considered a major deterrent to the formation of self-employment (Taylor, 2001). This paper analyzes different types of self-employment in urban China, and examines how a major policy change, China's 2007 Property Rights Law that improved the access to credit, affects the formation of different types of businesses. Utilize China Health and Nutrition Survey and Chinese Private Enterprises Survey data, and difference-in-difference models, I investigate whether the improvement of credit constraint is internalized by these two types of self-employed workers in their labor market decisions.

China provides an excellent laboratory for this analysis because of the presence of large heterogeneity among self-employed workers, the recent rise of the private sector, and the exogenous change in policy that relax the credit constraint. Official statistics show the number of self-employed workers in urban China increased from 0.15 million in 1978, the beginning of the economic transition, to 61.42 million in 2013. Some of the newly self-employed are true entrepreneurs seeking opportunities in the transition economy, but others are forced into self-employment out of necessity. There are also striking regional variations in the level of economics development and self-employment rate. These variations, both over time and across regions, enable me to estimate the impact of policy interventions statistically.

In this paper, I first establish a model of employment decision that incorporates credit constraint and employee hiring choices. At equilibrium, individuals who choose self-employment are separated into two distinctive types: the ones with highest ability and the ones with the lowest. These two types of self-employed workers correspond to the notion of opportunity entrepreneurs and necessity entrepreneurs. My model also predicts that high ability entrepreneurs hire labor while low ability entrepreneurs do not.[1] Thus, empirically, I separate urban self-employed workers into two types — *self-employed own-account workers* and *self-employed owner-managers with paid employee(s)*. This separation criteria is also used in Earle and Sakova (2000), showing that in six transition economies, owner-managers with employees tend to be true entrepreneurs, but at least some own-account workers are somehow constrained and are involuntarily self-employed.

I then examine the impact of credit constraints on the formation of different types of private businesses by looking at the differences before and after the passage of the 2007 Property Rights Law. For small businesses in China, the type of assets which can be used as collateral is heavily restricted. This constraint was relaxed in 2007, when China's Property Rights Law was enacted. The law expands the scope of collateral that can be used by borrowers to secure loans. To obtain the causal effect of this relaxation of credit constraints,

[1]Models by Kihlstrom and Laffont (1979) and Van Praag and Cramer (2001), among the others, also link entrepreneurship to labor demand, predicting a positive relationship between entrepreneurial ability and the number of employees they hire.

I employ a difference-in-differences strategy, comparing the differences in self-employment rate before and after 2007, and for provinces with different marketization levels. I use the Index of Marketization by Fan et al. (2003) as the measure of the marketization level. I found that the relaxation of credit constraint significantly increases the formation of opportunity entrepreneurs, but has little impact on necessity entrepreneurs.

My study has two distinctive contributions. First, it contributes to the credit constraint literature that examines the individual-level decision to enter into entrepreneurship. Carpenter and Petersen (2002) and Banerjee and Duflo (2014), among the others, found evidence of severe credit constraint for private businesses. I took advantage of a nationwide policy change to provide more empirical evidence on whether the improvement of credit constraint through collateral is internalized by self-employed workers.

Second, I address the heterogeneity among self-employed workers. More specifically, I separate them into necessity entrepreneurs and opportunity entrepreneurs, and formalize it in a theoretical model. I then examine the heterogeneous effect of credit constraint on them. Necessity entrepreneurs prevail in less developed countries, where self-employment rate is higher than developed countries (Parker, 2004). La Porta and Shleifer (2008) presented evidence that many of the self-employed in developing countries run low-productivity, commonplace businesses. Mohapatra et al. (2007) evaluated the role of self-employment in China's rural economy, finding that self-employers with low level of capital invested tend to be necessity entrepreneurs. Tokman (1992) also showed that in Latin America, self-employment is primarily a refuge for people that are excluded from formal labor markets. In my paper, I found that credit constraint is a major deterrent for opportunity entrepreneurs, but not so much for necessity entrepreneurs.

The rest of this paper is organized as follows: Section 4.2 introduces the institutional background and the passage of the 2007 Property Rights. Section 4.3 constructs a model and present propositions on the impact of credit constraint. Section 4.4 discusses data source and presents the basic patterns in the study population. Section 4.4 also analyzes the characteristics of the two types of self-employment. Section 4.5 presents the empirical strategies and 4.6 the corresponding results. Section 4.7 discusses the results and concludes.

4.2 INSTITUTIONAL BACKGROUND

Credit constraint and the lack of access to formal sources of external finance are often considered as a major deterrent to business formation around the world (?). In China, one peculiar problem for small businesses is that the type of assets which can be used as collaterals is heavily restricted. According to Han (2007), while 70% of small business financing was secured by movable property in the U.S., the figure for China was less than 15%. World Bank (2007) estimated that small and medium sized enterprises and farmers in China have approximately $2 trillion in "dead capital" – assets that can not be used to generate loans to fund investment and growth.

Rules regarding collateral requirements in China are established in 1995 through the Collateral Law. Twenty years later, following legislative debates that had lasted for over a decade, China passed its first Property Rights Law. This Property Rights law was passed on March 16, 2007 and went into effect in October 1, 2007.[2] The 2007 Property Law emphasizes, for the first time, the equal protection of state, collective and individual property. Moreover, it reforms the collateral requirements, and relaxes the credit constraint for entrepreneurs in three major ways.

First, the law expands the scope of collaterals that can be used by borrowers to secure loans (Marechal et al., 2009; Berkowitz et al., 2015; Du, 2010). It enables borrowers to use a broader level of movable collateral, such as present and future-acquired equipment inventory, accounts receivable and a combination of assets (Clause 181). For example, in the 1995 Collateral Law, only properties proved by law or regulation can be used as collaterals; correspondingly, the 2007 Property Rights Law states in Clause 180 that "properties other than those that shall not be mortgaged according to any law or administrative regulation" can be used as collateral. Second, the law improves the process and simplifies the formalities required to obtain a loan. Third, the law reduces the liquidation cost for banks which incentivizes them to lend (Liu, 2008). For example, previously, when the obligor fails to

[2]On December 29, 2006, the 25th session of the 10th standing committee of the National People's Congress accepted a draft of the Property Rights Law of the People's Republic of China. Once the Standing Committee approved the draft of the law, its formal passage was almost a done deal. Berkowitz et al. (2015) showed that the pass of this law came as a surprise, because at the time it was uncertain whether the law would go forward to the full session of the National People's Congress.

pay his/her due debts and the mortgagee and the mortgagor fail to conclude an agreement on the means of realizing the right to mortgage, the mortgagee need to go through lengthy law suits. However, Clause 195 of the Property Law states that in these scenarios, the mortgagee may request the people's court to auction or sell off the property under mortgage. This improvement is recognized by the World Bank: China moved up 20 places (out of 181 economies) on the "Getting Credit – Legal Rights" ranking in its *Doing Business* Index in 2008.

4.3 THEORETICAL FRAMEWORK

I develop a static partial equilibrium model of employment choice that incorporates both liquidity constraint and employee hiring choices. I then visualize the equilibrium of this model by simulating it in Figure 4.1. The model shows that individuals who choose self-employment are separated into two distinctive types: the ones with highest ability (opportunity entrepreneurs) and the ones with the lowest (necessity entrepreneurs). The simulation shows that the relaxation of credit constraint promote entrepreneurship, however, this effect is significant for the opportunity entrepreneurs but trivial on necessity entrepreneurs.

In this model, individuals choose whether to be self-employed or work for a wage. Suppose individuals differ in ability θ and wealth z, and they know their own ability before committing to their occupational choices. Both θ and z are normalized to range between 0 and 1. I assume that wage w and capital rent (i.e. one plus interest rate) r are exogenous.

4.3.0.1 Optimization for Self-Employers Self-employers choose the amount of capital invested in the business and the number of employees to maximize their profit, given the level of r and w. I assume the production function is in the form of $g(\theta)f(k,n) = e^{\theta}k^{\alpha}(n+\epsilon)^{\beta}$. Individual's ability θ affects productivity exponentially because of the "superstar" theory – high ability entrepreneurs earn exponentially more than low ability entrepreneurs. k is the amount of capital used and n is the number of "skill unit" hired. Entrepreneurs themselves count as ϵ unit of labor, so the total amount of labor is $n+\epsilon$. Following Buera et al. (2013), α

and β are the elasticities of output with respect to capital and labor and $\alpha + \theta < 1$, implying diminishing returns to scale in variable factors at establishment level.

The self-employed workers face capital constraint similar as Evans and Jovanovic (1989). Each person can borrow up to an amount that is proportional to his/her wealth: $(\lambda - 1)z$, so the most a person can invest in the business is $z + (\lambda - 1)z = \lambda z$. The 2007 Property Rights law broadened the extent of collateral and improved the access to credit, thus increased λ.

To summarize, the optimization problem for self-employed workers is as follow:

$$\max_{k,n} \; e^{\theta} k^{\alpha} (n + \epsilon)^{\beta} + r(z - k) - wn \tag{4.1}$$

$$s.t. \; 0 \leq k \leq \lambda z, n \geq 0 \tag{4.2}$$

Note that $k = 0$ will never be the optimal choice for the entrepreneurs, so the constraint $k \geq 0$ will never bind. I can then derive the solution (k^*, n^*) to this maximizing problem in four cases:

1. None of the constraints are binding, i.e. $k^* < \lambda z, n^* > 0$:

$$\begin{cases} k^* = (\frac{\alpha^{1-\beta}\beta^{\beta}}{r^{1-\beta}w^{\beta}} e^{\theta})^{\frac{1}{1-\alpha-\beta}} \\ n^* = (\frac{\beta^{1-\alpha}\alpha^{\alpha}}{w^{1-\alpha}r^{\alpha}} e^{\theta})^{\frac{1}{1-\alpha-\beta}} - \epsilon \end{cases} \tag{4.3}$$

   Solving the non-binding condition $k^* < \lambda z$ and $n^* > 0$ gives us the following inequalities:

$$k^* < \lambda z \Leftrightarrow \theta < (1 - \alpha - \beta) ln(\lambda z) - ln(\frac{\alpha^{1-\beta}\beta^{\beta}}{r^{1-\beta}w^{\beta}}) \tag{4.4}$$

$$n^* > 0 \Leftrightarrow \theta > (1 - \alpha - \beta) ln(\epsilon) - ln(\frac{\beta^{1-\alpha}\alpha^{\alpha}}{w^{1-\alpha}r^{\alpha}}) \tag{4.5}$$

   The intuition behind these two inequalities is that entrepreneurs will only hire employees when their own ability θ is larger than a threshold value, and they will be unconstrained in capital if their ability is relatively small compared to their wealth.

2. The constraint on labor binds but the constraint on capital does not:

$$\begin{cases} k^* = (\frac{\alpha\epsilon^{\beta}}{r} e^{\theta})^{\frac{1}{1-\alpha}} \\ n^* = 0 \end{cases} \tag{4.6}$$

3. The constraint on capital binds but the constraint on labor does not:

$$\begin{cases} k^* = \lambda z \\ n^* = (\frac{(\lambda z)^\alpha \beta}{w} e^\theta)^{\frac{1}{1-\beta}} - \epsilon \end{cases} \tag{4.7}$$

4. Both constraints are binding:

$$\begin{cases} k^* = \lambda z \\ n^* = 0 \end{cases} \tag{4.8}$$

Plugging (k^*, n^*) into the income structure equation, we can obtain the optimized income for self-employed workers Y^*_{SE}:

$$Y^*_{SE} = \begin{cases} (1-\alpha-\beta)(\frac{\alpha^\alpha \beta^\beta}{r^\alpha w^\beta} e^\theta)^{\frac{1}{1-\alpha-\beta}} + w\epsilon + rz & \text{if } \theta \text{ satisfies eq 4.4 and 4.5} \\ (1-\alpha)(\frac{\alpha}{r})^{\frac{\alpha}{1-\alpha}} (\epsilon^\beta e^\theta)^{\frac{1}{1-\alpha}} + rz & \text{if } \theta \text{ satisfies eq 4.4 but not 4.5} \\ (1-\beta)(\frac{\beta}{w})^{\frac{\beta}{1-\beta}} ((\lambda z)^\alpha e^\theta)^{\frac{1}{1-\beta}} + r(1-\lambda)z + w\epsilon & \text{if } \theta \text{ satisfies eq 4.5 but not 4.4} \\ (\lambda z)^\alpha \epsilon^\beta e^\theta + r(1-\lambda)z & \text{if otherwise} \end{cases} \tag{4.9}$$

4.3.0.2 Income for Wage-Earners For wage-earners, they get paid w based on their skill as workers, and can get capital return on their endowment z. I assume that individuals' skill as wage workers is the same as their ability θ. There is a probability P that they can not find a job. I assume that this probability is a power function of their ability θ. This setting allows me to incorporate the existence of "necessity entrepreneurs". Some individuals will gladly work for wages but could not find jobs with payment matching their expectation, so they turn to self-employment out of necessity. The income structure for wage earners Y_W is as below:

$$Y_W = P(\theta)\theta w + rz = \theta^{\tau+1} w + rz \tag{4.10}$$

4.3.0.3 Optimal Occupational Choices Knowing their entrepreneurial income, individuals select into self-employment by comparing their optimal income as self-employed workers Y_{SE}^* and as wage-earners Y_W. An individual will only choose to start a business if and only if his/her expected income from doing so exceeds that from wage work:

$$e^{\theta}(k^*)^{\alpha}(n^* + \epsilon)^{\beta} + r(z - k^*) - wn^* \geq \theta^{\tau+1}w + rz \tag{4.11}$$

I analyze inequality 4.11 in four scenarios based on whether the capital constraint and hiring constraint are binding. The occupational choices for individuals with ability θ and wealth z are as below:

1. For individuals who do not subject to capital constraints and have high ability above the threshold to hire employee(s), they choose self-employment if and only if the following equations are satisfied:

$$\begin{cases} (1-\alpha-\beta)(\frac{\alpha^{\alpha}\beta^{\beta}}{r^{\alpha}w^{\beta}}e^{\theta})^{\frac{1}{1-\alpha-\beta}} + w\epsilon \geq \theta^{\tau+1}w \\ \theta \geq (1-\alpha-\beta)ln(\epsilon) - ln(\frac{\beta^{1-\alpha}\alpha^{\alpha}}{w^{1-\alpha}r^{\alpha}}) \\ \theta \leq (1-\alpha-\beta)ln(\lambda z) - ln(\frac{\alpha^{1-\beta}\beta^{\beta}}{r^{1-\beta}w^{\beta}}) \end{cases} \tag{4.12}$$

2. For those who are unconstrained in capital but have low ability below the threshold to hire employee(s), they choose self-employment if and only if the following equations are satisfied:

$$\begin{cases} (1-\alpha)(\frac{\alpha}{r})^{\frac{\alpha}{1-\alpha}}(\epsilon^{\beta}e^{\theta})^{\frac{1}{1-\alpha}} \geq \theta^{\tau+1}w \\ \theta < (1-\alpha-\beta)ln(\epsilon) - ln(\frac{\beta^{1-\alpha}\alpha^{\alpha}}{w^{1-\alpha}r^{\alpha}}) \\ \theta \leq (1-\alpha-\beta)ln(\lambda z) - ln(\frac{\alpha^{1-\beta}\beta^{\beta}}{r^{1-\beta}w^{\beta}}) \end{cases} \tag{4.13}$$

3. For those who are constrained in capital but have high ability above the threshold to hire employee(s), they choose self-employment if and only if the following equations are satisfied:

$$\begin{cases} (1-\beta)(\frac{\beta}{w})^{\frac{\beta}{1-\beta}}((\lambda z)^{\alpha}e^{\theta})^{\frac{1}{1-\beta}} - r\lambda z + w\epsilon \geq \theta^{\tau+1}w \\ \theta \geq (1-\alpha-\beta)ln(\epsilon) - ln(\frac{\beta^{1-\alpha}\alpha^{\alpha}}{w^{1-\alpha}r^{\alpha}}) \\ \theta > (1-\alpha-\beta)ln(\lambda z) - ln(\frac{\alpha^{1-\beta}\beta^{\beta}}{r^{1-\beta}w^{\beta}}) \end{cases} \tag{4.14}$$

4. For those who are constrained in capital and have low ability below the threshold to hire employee(s), they choose self-employment if and only if the following equations are satisfied:

$$\begin{cases} (\lambda z)^{\alpha}\epsilon^{\beta}e^{\theta} - r\lambda z \geq \theta^{\tau+1}w \\ \theta < (1-\alpha-\beta)ln(\epsilon) - ln(\frac{\beta^{1-\alpha}\alpha^{\alpha}}{w^{1-\alpha}r^{\alpha}}) \\ \theta > (1-\alpha-\beta)ln(\lambda z) - ln(\frac{\alpha^{1-\beta}\beta^{\beta}}{r^{1-\beta}w^{\beta}}) \end{cases} \tag{4.15}$$

I simulate all four scenarios in a (θ, z) space, and plot them in Figure 4.1, which shows the nature of selection.[3] Two yellow parts represent individuals who chose self-employment and subject to no capital constraint. Two blue parts are self-employed workers who are capital constrained. The remaining white parts are wage workers.

When there is no capital constraint, i.e. $\lambda = \infty$, there are only unconstrained individuals. The solution simplified to $\theta \in (-\infty, \bar{\theta_1}] \cup [\bar{\theta_2}, \infty)$, where $\bar{\theta_1}$ and $\bar{\theta_2}$ are the value of the two solid red vertical lines in Figure 4.1. Thus, anyone with ability above $\bar{\theta_2}$ or below $\bar{\theta_1}$ will choose self-employment, while others will chose to work for wage.

Although both high ability people and low ability people are potential self-employed workers, they are different. The would-be self-employed individuals on the upper end (right side) of the ability distribution hire employees, while the ones on the lower end (left side) of the ability distribution do not hire employees. The results from this model is consistent with models by Kihlstrom and Laffont (1979) and Van Praag and Cramer (2001), among the others, which link entrepreneurship to labor demand, predicting a positive relationship between entrepreneurial ability and the number of employees they hire. The self-employed workers with low ability choose self-employment because they have low probability of getting a wage job, resulting in low expected wage earning. The self-employed workers with high ability choose self-employment because they have high potential entrepreneurial earning. Correspondingly, they are "necessity entrepreneurs" and "opportunity entrepreneurs". This separates my work from other labor market choice models.

[3] I assigned values to the parameters following Buera et al. (2013): $\alpha = 0.2$, $\beta = 0.59$, $r = 1.07$, $w = 1$, $\tau = 0.3$, $\lambda = 2$

Now consider a relaxing in credit constraint that λ increased by $\Delta\lambda$. The blue lines in Figure 4.2 shows the new occupational choices after this change. Some self-employers' constraint become not binding; some wage workers choose to switch to self-employment because they can borrow more and invest more which increase their potential entrepreneurial income. The green shades mark the new self-employed workers after the relaxing of credit constraint: the shade on the right corresponding to "opportunity entrepreneurs" is substantial, while the shade on the left corresponding to "necessity entrepreneurs" is minimal. This is because necessity entrepreneurs usually have very low ability so the amount of optimal capital they required is limited.

Based on Figure 4.2, I have the following three propositions on the changes in self-employment after credit constraint changes from λ to $\lambda + \Delta\lambda$:

Proposition 1. *For the same $\Delta\lambda$, the lower the initial λ, the higher the increase in the self-employment rate.*

Proposition 2. *For the same λ, the higher the $\Delta\lambda$, the higher the increase in the self-employment rate.*

Proposition 3. *When λ increases, the increase in the rate of owner-manager is much larger than the increase in the rate of own-account workers.*

I test these three propositions by taking advantage of the passage of the 2007 Property Rights in China, which broadened the extent of collateral which relaxes the credit constraint, corresponding to an increase in λ in my model. I predict that a relaxing in credit constraint will increase self-employment propensity, but it will affect opportunity entrepreneurs more than necessity entrepreneurs.

4.4 DATA

4.4.1 Data Source

I utilize several different datasets: the China Health and Nutrition Survey (CHNS), the Chinese Private Enterprises Survey (CPES), the National Economic Research Institute Marketization Index (NERI Index), and data from the China Statistics Bureau.

I use five waves of the CHNS data: 2000, 2004, 2006, 2009, and 2011. CHNS is a nationally representative dataset that covers both urban and rural areas, collected by the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention (Zhang et al., 2014). I exclude data from rural sites since our area of focus is the urban labor market. Before 2011, this survey was conducted in nine provinces including Liaoning, Heilongjiang, Jiangsu, Shandong, Henan, Hubei, Hunan, Guangxi, and Guizhou. In the 2011 wave, Beijing, Shanghai, and Chongqing were added. I exclude these late additions from our analysis. CHNS is a panel dataset and the same individuals were interviewed in each wave. The annual attrition rate is around 30% and new individuals were added to maintain the representativeness of the sample (See Appendix C for a detailed discussion on attrition and data representativeness).

Chinese Private Enterprises Survey (CPES) is a firm level dataset, conducted every two year from 2000 to 2012 by the Privately Owned Enterprises Research Project Team. All the subjects are owner-managers, and the dataset contains detailed information both on the firm and on the personal background of the owner-manager.

The National Economic Research Institute Marketization Index is an extensively used database containing the institutional index developed by the Nevin Economic Research Institute to reflect the regional institutional environment (Fan et al., 2003, 2011). The index is constructed annually for each province based on five aspects: relationship between government and market, development of non-state sector, development of commodity market, development of factor market, and development of market intermediaries and legal environment. The index value ranges between 0 and 28, higher representing more market-oriented.

4.4.2 Two Different Types of Self-Employment

My model shows that one main distinction between the two types of self-employment is the number of employees they hire. Thus, I differentiate urban self-employed workers into two types based on whether they have paid employees: *self-employed own-account workers* and *self-employed owner-managers*. The former work on their own, while the latter have at least one paid employees. Figure 4.3 presents the time trend of the self-employment rate in China between 1989 and 2011 using all waves of CHNS data. Here, the self-employment rate is defined as the proportion of these two types of self-employed individuals among all study subjects in each wave. The percentage of self-employed owner-managers increased over time, while the percentage of self-employed own-account workers reached a peak around 1997 then started to decline after 2000.

I use education level as a crude indicator for individual's ability, and plot the time trend of average years of schooling for owner-manager and own-account workers in Figure 4.4, along with their 95% confidence interval. Consistent with my model, there is a persistent 3-year gap in years of schooling among them, with own-account workers being less educated. More analysis in Appendix E also shows that own-account workers are often not from well-educated families and with low income. Own-account workers earned lower salary when they were wage workers compared to the remaining wage earners. After their transition into self-employment, their income decreased even further. For owner-mangers who were salary workers in the past, there is no difference in their wage compared to the remaining wage earners. After the transition, their income increased significantly. All evidence I found are consistent with my theory and my model, suggesting that self-employed own-account workers are "necessity entrepreneurs", while self-employed owner-manger's entrepreneurship is an unconstrained choice.

Empirically, it is also natural to use the number of employees as the separation criteria for different types of self-employment in China, since it is an important feature in the separation of different enterprise categories: *Getihu* are private businesses that are registered with no more than seven people hired as employees; *Siyingqiye* are those with more than seven employees. Besides the difference in the maximum number of employees, Getihu also receives more lenient policy regarding the assets they could use for running the firms (Zhang and Van Stel, 2016).

This separation criterion is also used by Earle and Sakova (2000) in their analysis on six former Soviet countries. They found that the own-account workers tend to be necessity entrepreneurs which characterized by negative traits, like low education, low household income and low family income. I can not deny that there might be some true entrepreneurs with no employees, but my analysis provided in this paper shows that, this crude separation already sheds light on the distinctive characteristics of different kinds of self-employment.[4]

4.4.3 Covariates

Guided by the literature on the empirical studies of entrepreneurship (Le (1999) etc.), I control for various individual characteristics and provincial characteristics. Individual characteristics include gender, age (and age squared), education (measured by years of schooling), marriage status (1 indicating married and 0 otherwise), individual annual income, and whether the spouse works. I also test alternative indicators, such as using the level of completed education instead of using years of schooling, and the results are similar. Provincial characteristics include regional gross domestic production (1000 USD), province population, number of industrial enterprises above designated size, disposable income of urban residents, total value of import and export from the region (1000 USD), and value added of the manufacturing industry.

Summary statistics for different samples in CHNS are shown in Table 4.1. Consistent with Figure 4.4 and Appendix E, own-account workers are the least educated and earn the least amount of money.

[4] This is not the only separation criterion used in the literature. Levine and Rubinstein (2015) disaggregate the self-employed in the U.S. into incorporated and unincorporated. Their results suggest that the incorporated ones are true entrepreneurs with distinctive cognitive and noncognitive traits, and they earn much more per hour than their salaried and unincorporated counterparts. However, their criterion does not suit developing countries. In the U.S., the self-employment rate is about 11%, a third of which are incorporated. In the developing countries, the vast majority of self-employment are one man firms and the number of incorporated ones are very small. What's more, the legal system in most developing countries are less complete, making it difficult to interpret the implication of being incorporated and unincorporated. Mohapatra et al. (2007) evaluate the role of self-employment in China's rural economy, separating the self-employed into two types based on the level of capital invested in the self-employed business. They find that the high productivity type is a sign of development while the low productivity type is a sign of distress. This criterion is hard to apply here because of data restriction.

4.5 EMPIRICAL STRATEGY

4.5.1 Impact on Self-Employment Propensity

I intend to show the causal effect of credit constraint on self-employment propensity after the passage of the 2007 Property Rights Law. Since different provinces can vary in their original status of the financial market and the execution of the law, the degree of relaxation in credit constraint due to the passage of the law can be different. Thus, I utilize a difference-in-differences strategy by taking advantage of the provincial variations, measured by the NERI Marketization Index.

I focus on three different dimensions of the NERI index that are directly related to the collateral channel: credit market marketization level, financial marketization level, and the protection of producers. Credit marketization level represents the share of loans to non-state-owned businesses. Financial marketization level is a weighted average of the credit marketization level and the competitiveness of the financial market. Protection of the producers represents whether the local legal system execute the law justly.

Figure 4.5 shows the NERI marketization index on financial marketization level in 2007 and 2008 for the provinces in my sample. Using the model presented in Section 4.3, I predict that: (1) Controlling for the change in NERI index, self-employment rate is higher for provinces with tighter initial constraint in 2007, i.e. lower initial value in 2007; (2) Controlling for 2007 NERI index, self-employment rate is higher for provinces with larger increase in the NERI index; (3) The impact on owner-managers is larger than own-account workers.

Empirically, I estimate the following two models: one with DID interaction terms of post 2007 indicator and the NERI index, and another with the interaction terms of post 2007 indicator and the change in the NERI index $\Delta NERI$:

$$\begin{aligned} OwnAccount_{ict} = \beta_{10} + \beta_{11} Post_t + \beta_{12} NERI_c * Post_t + \beta_{13} \Delta NERI_c + \\ \beta_{14} X_{ict} + \beta_{15} Province_{ct} + \beta_{16} Year_t + \beta_{17} City_{it} + \epsilon_1 \\ OwnerManager_{ict} = \beta_{20} + \beta_{21} Post_t + \beta_{22} NERI_c * Post_t + \beta_{23} \Delta NERI_c + \\ \beta_{24} X_{ict} + \beta_{25} Province_{ct} + \beta_{26} Year_t + \beta_{27} City_{it} + \epsilon_2 \end{aligned} \tag{4.16}$$

$$
\begin{aligned}
OwnAccount_{ict} = \beta_{30} + \beta_{31} Post_t + \beta_{32} \Delta NERI_c * Post_t + \beta_{33} NERI_c + \\
\beta_{34} X_{ict} + \beta_{35} Province_{ct} + \beta_{36} Year_t + \beta_{37} City_{it} + \epsilon_3 \\
OwnerManager_{ict} = \beta_{40} + \beta_{41} Post_t + \beta_{42} \Delta NERI_c * Post_t + \beta_{43} NERI_c + \\
\beta_{44} X_{ict} + \beta_{45} Province_{ct} + \beta_{46} Year_t + \beta_{47} City_{it} + \epsilon_4
\end{aligned} \tag{4.17}
$$

where $OwnAccount_{ict}$ and $OwnerManager_{ict}$ are indicators for individual i being own-account workers or owner-manager in province c at year t. $Post$ indicates after China's 2007 Property Rights Law. $NERI_c$ is the marketization index for province c in 2007 and $\Delta NERI_c$ is the change in NERI index from 2007 to 2009 for proince c. X_{ict} and $Province_{ct}$ are the individual characteristics and province characteristics introduced in Section **??**. Year fixed effect and city fixed effect are also included. Standard errors are clustered at the provincial level.

If my model predictions are correct, β_{22} should be negative and β_{42} should be positive, while β_{12} and β_{32} might be statistically insignificant. For necessity entrepreneurs, the policy has less impact, thus I hypothesize β_{12} to be insignificant.

4.5.2 Impact through Collateral Channel

In order to verify that the effect indeed comes from the relaxation of credit constraint, I use CPES data to create a binary index $BankBorrow$ that represents this channel. Subjects were asked whether they borrowed from the banks when starting their businesses. If the answer is "yes" for owner-manager i who founded his/her business in year t in province c, then $BankBorrow_{ict} = 1$; and 0 otherwise. I use $BankBorrow$ as the outcome variable and run a similar difference-in-differences model as below:

$$
\begin{aligned}
BankBorrow_{ict} = \gamma_{10} + \gamma_{11} Post_t + \gamma_{12} NERI_c * Post_t + \gamma_{13} \Delta NERI_c + \\
\gamma_{14} X_{ict} + \gamma_{15} Province_{ct} + \gamma_{16} Year_t + \gamma_{17} City_{it} + \epsilon_1 \\
BankBorrow_{ict} = \gamma_{20} + \gamma_{21} Post_t + \gamma_{22} \Delta NERI_c * Post_t + \gamma_{23} NERI_c + \\
\gamma_{24} X_{ict} + \gamma_{25} Province_{ct} + \gamma_{26} Year_t + \gamma_{27} City_{it} + \epsilon_2
\end{aligned} \tag{4.18}
$$

If the 2007 Property Law indeed relaxed the credit constraint and encouraged more entrepreneurs to borrow from the banks, provinces with lower initial NERI value and with larger increase in the NERI index should observe more increase in bank loans. Thus, γ_{12} should be negative and γ_{22} should be positive.

4.6 RESULTS

I utilized a difference-in-differences strategy in order to show the causal effect of credit constraint on self-employment propensity. Results for this model are reported in Table 4.2. Three columns correspond to the three NERI indexes. In Panel A, the outcome variable is $OwnerManager$, the indicator for being owner-manager. The coefficients for the interaction terms $Post * NERI$ are all significantly negative, ranging from 1.23% to 2.24%. The coefficients for the interaction terms $Post * \Delta NERI$, however, are all significantly positive. This implies that the larger the relaxation in credit constraint, the larger the increase in self-employment propensity for opportunity entrepreneurs. The results are consistent with Proposition 1 and Proposition 2. In Panel B, the coefficients for all the interaction terms are not significant, suggesting that the law has little impact on own-account worker, consistent with Proposition 3.

To verify that the channel is through the relaxation of credit constraint, another difference -in-differences model is used and the results are reported in Table 4.3. As predict, the coefficients are all negative on the interaction terms $Post * NERI$ and positive on $Post * \Delta NERI$. This implies that the would-be owner-managers living in provinces with initial worse credit market or provinces experienced larger expansion in credit market, borrowed more from the banks after the 2007 Property Rights Law.

All the evidence suggest that the 2007 Property Rights Law improved the credit constraint, which promotes entrepreneurship represented by the amount of owner-manger in the economy, but it has little impact on necessity entrepreneurs.

4.7 DISCUSSION AND CONCLUSION

This paper builds upon unique Chinese circumstances to explore the prevalence and origins of different types of entrepreneurs, and how limited access to credit might impair them in different ways. I distinguish and analyze two different types of self-employment in urban China—opportunity entrepreneurship and necessity entrepreneurship. Opportunity entrepreneurs are owner-managers who exploit business opportunities, hire workers, and promote overall economic activity. Necessity entrepreneurs are individuals who turn to self-employment as a last resort. I construct a structural model that contains both liquidity constraints and employee hiring choices to make predictions on the corresponding self-employment rates after a relaxation of credit constraints. I test the predictions empirically, finding that the 2007 Property Rights Law promotes entrepreneurship as measured by the number of owner-mangers in the economy, but has little impact on necessity entrepreneurs. Studying the formerly neglected distinction between "true entrepreneurs" and "entrepreneurs by default" opens the door to more insightful analyses of enterprise formation that are valuable for China and many other countries.

Figure 4.1: Selection into Self-Employment

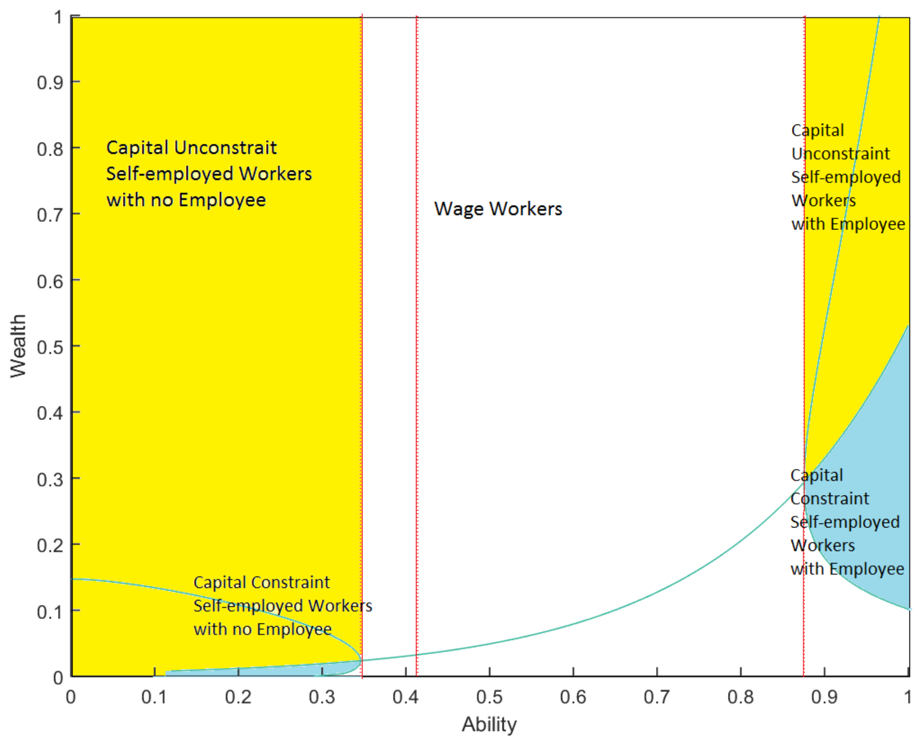
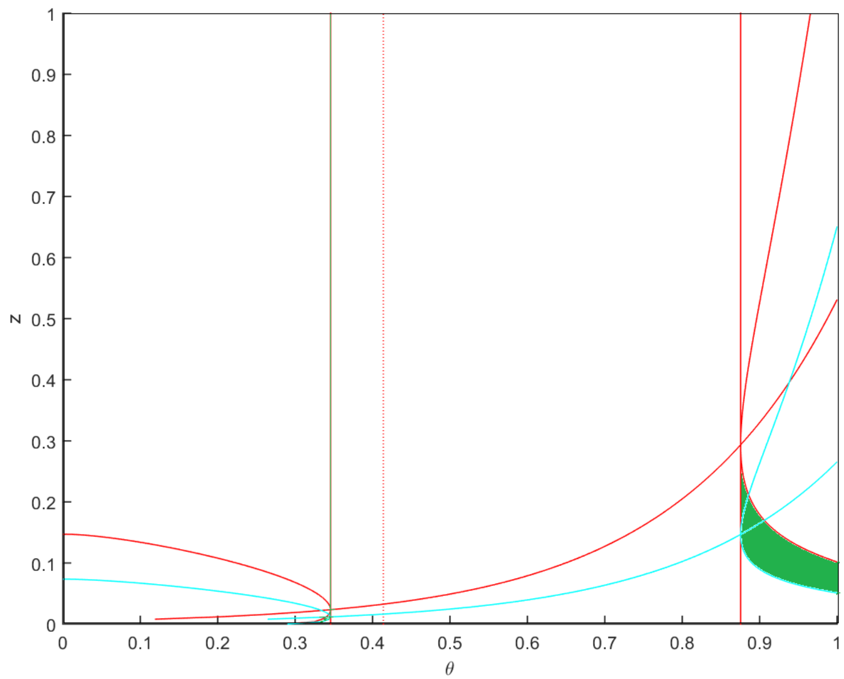


Figure 4.2: Change in Credit Constraint



Note: The red lines represent the occupational choices under original credit constraint, and the blue lines represent after relaxing the credit constraint. The green shades mark the new self-employed individual who switch from wage workers.

Figure 4.3: Self-Employment Rate (%): 1989-2011

Note: Self-employment rate for own-account worker is defined as the proportion of own-account workers among all study subjects in each wave. Similarly, self-employment rate for owner-manager is the proportion of owner-managers among all study subjects in each wave.

Figure 4.4: Gaps in Average Years of Schooling: 1989-2011

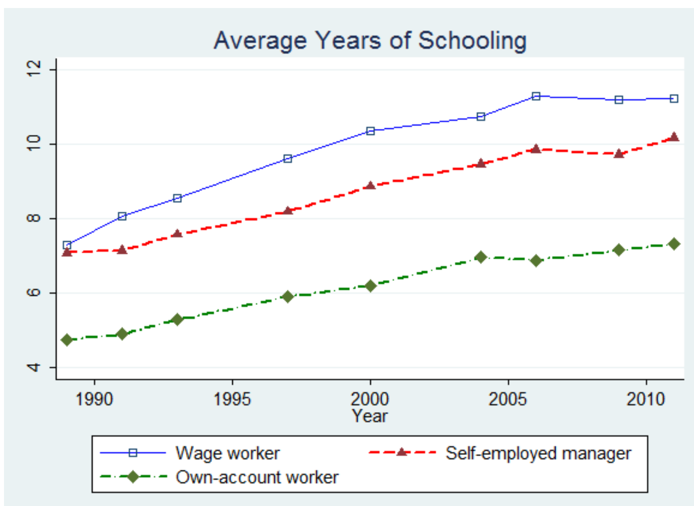


Table 4.1: Summary Characteristics (1989-2011 Combined)

	All Sample	Salary Worker	Own-account	Manager
Female Ratio (%)	47.12	45.73	51.08	38.53
	(49.92)	(49.82)	(49.99)	(48.71)
Age (year)	40.29	39.97	41.14	38.92
	(13.63)	(13.50)	(14.13)	(11.00)
Education (year)	8.342	9.430	5.877	9.022
	(4.365)	(4.133)	(3.920)	(3.449)
Marriage Rate (%)	79.76	78.44	82.25	84.76
	(40.18)	(41.12)	(38.21)	(35.97)
Urban Hukou (%)	64.01	83.97	20.21	66.37
	(48.00)	(36.69)	(40.16)	(47.29)
Change Job (%)	12.49	15.23	5.509	21.28
	(33.06)	(35.93)	(22.82)	(40.98)
Annual Income (RMB)	8063.2	9044.4	4904.3	17717.4
	(15570.0)	(16123.1)	(10815.3)	(30849.5)
Annual HH Income (RMB)	20432.9	22844.9	13506.4	35638.0
	(29733.2)	(31538.2)	(20601.4)	(46782.1)
Work Hour (hour)	2205.3	2179.1	2259.8	2623.1
	(616.2)	(468.3)	(1080.3)	(793.8)
Spouse Has Job (%)	54.50	50.54	63.11	57.71
	(49.80)	(50.00)	(48.25)	(49.45)
Observations	20170	13568	6018	584

Notes: Mean values are reported, and standard deviations are in parenthesis. All samples are for adult residents living in the urban areas.

Figure 4.5: NERI Financial Marketization Index in 2007 and 2008

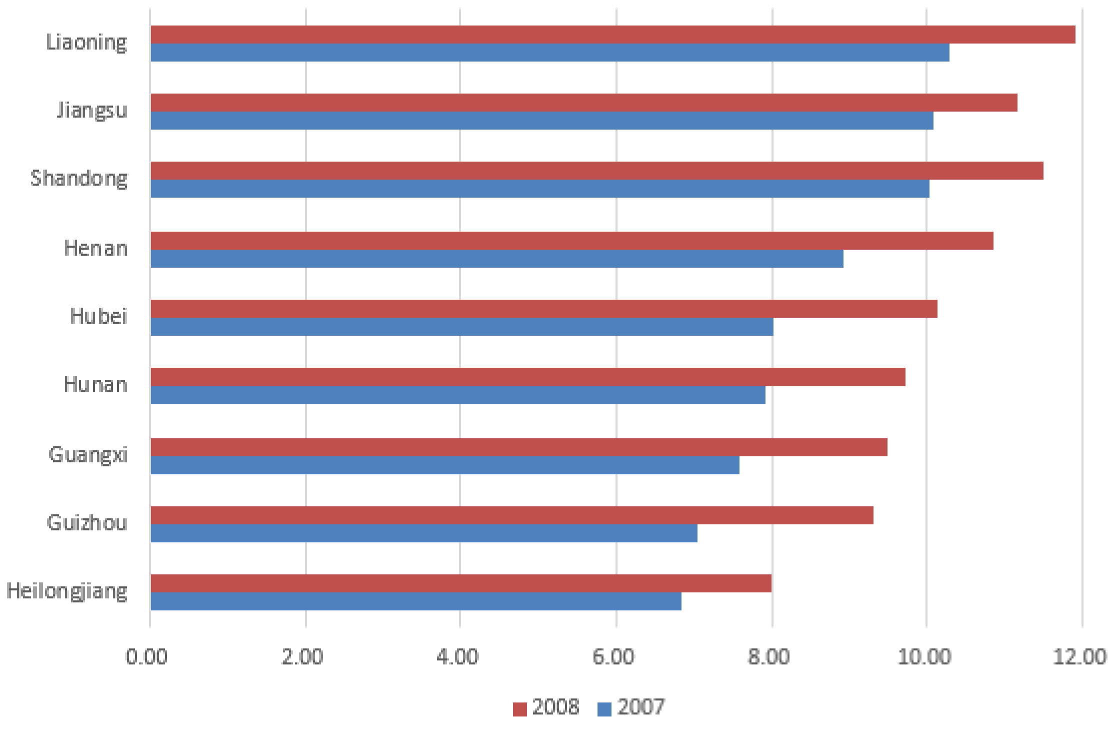


Table 4.2: Impact on Self-Employment Propensity

	(1)	(2)	(3)
NERI Type	Financial Market	Credit Marketization	Producer Protection
Panel A: Outcome is being owner-manager			
Post*$NERI$	-0.0209*	-0.0224**	-0.0123*
	(0.0104)	(0.00933)	(0.00619)
Post*$\Delta NERI$	0.0345**	0.0439***	0.0302***
	(0.0126)	(0.0129)	(0.00649)
Panel B: Outcome is being own-account workers			
Post*$NERI$	0.0332	-0.00608	0.0161
	(0.0296)	(0.0298)	(0.0110)
Post*$\Delta NERI$	0.0107	-0.0283	-0.0117
	(0.0613)	(0.0452)	(0.0236)
Observations	7,567	7,567	7,567

Notes: Each cell is a different regression. Standard errors are clustered at provincial level: *** p<0.01, ** p<0.05, * p<0.1

Table 4.3: Impact through Collateral Channel

	(1)	(2)	(3)
Post*$NERI$	-0.00828	-0.0144**	-0.0213**
	(0.0117)	(0.00492)	(0.00677)
Post*$\Delta NERI$	0.104***	0.0369**	0.0171*
	(0.0251)	(0.0113)	(0.00873)
NERI Type	Financial Market	Credit Marketization	Producer Protection
Observations	1,493	1,493	1,493

Notes: The outcome is an index for whether borrowed from the banks when founding the business. Each cell is a different regression. Standard errors are clustered at provincial level: *** p<0.01, ** p<0.05, * p<0.1

5.0 BIBLIOGRAPHY

Acs, Z. (2006). How is Entrepreneurship Good for Economic Growth? *Innovations 1*(1), 97–107.

Aggarwal, M. and I. A. Khan (2006). Hypertensive Crisis: Hypertensive Emergencies and Urgencies. *Cardiology Clinics 24*(1), 135–146.

Ai, C. and E. C. Norton (2003). Interaction Terms in Logit and Probit Models. *Economics letters 80*(1), 123–129.

Arora, S. (2001). Health, Human Productivity, and Long-Term Economic Growth. *The Journal of Economic History 61*(03), 699–749.

Bailey, J. B. (2013). Health Insurance and the Supply of Entrepreneurs: New Evidence from the Affordable Care Act's Dependent Coverage Mandate. *Working Paper*.

Baker, D. (2015). Job Lock and Employer-Provided Health Insurance: Evidence from the Literature. *AARP Public Policy Institute Working Paper*.

Banerjee, A. V. and E. Duflo (2014). Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program. *The Review of Economic Studies 81*(2), 572–607.

Bartel, A. and P. Taubman (1979). Health and Labor Market Success: The Role of Various Diseases. *The Review of Economics and Statistics*, 1–8.

Berkowitz, D., C. Lin, and Y. Ma (2015). Do Property Rights Matter? Evidence From a Property Law Enactment. *Journal of Financial Economics 116*(3), 583–593.

Bleakley, H. (2010). Health, Human Capital, and Development. *Annual Review of Economics 2*, 283.

Bound, J., T. Stinebrickner, and T. Waidmann (2010). Health, Economic Resources and the Work Decisions of Older Men. *Journal of Econometrics 156*(1), 106–129.

Buera, F. J., J. P. Kaboski, and Y. Shin (2013). The Macroeconomics of Microfinance. *Federal Reserve Bank of St. Louis Working Paper Series* (2013-034).

Carpenter, R. E. and B. C. Petersen (2002). Is the Growth of Small Firms Constrained by Internal Finance? *Review of Economics and statistics 84*(2), 298–309.

Chatterji, P., M. Alegria, and D. Takeuchi (2011). Psychiatric Disorders and Labor Market Outcomes: Evidence from the National Comorbidity Survey-Replication. *Journal of Health Economics 30*(5), 858–868.

Chobanian, A. V., G. L. Bakris, H. R. Black, W. C. Cushman, L. A. Green, J. L. Izzo Jr, D. W. Jones, B. J. Materson, S. Oparil, J. T. Wright Jr, et al. (2003). The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: The JNC 7 Report. *JAMA 289*(19), 2560–2571.

Chongqing Bureau of Statistics (2014). *Chongqing Statistics Yearbook [Chongqing Tongji Nianjian]*. China Statistics Press.

Chongqing Civil Affairs Bureau (2008). Chongqing urban and rural minimum living standards regulations [chongqing shi chengxiang zuidi shenghuo baozhang tiaoli].

Chongqing Human Resource Bureau (2012). Notice about Minimum Wage in Chongqing (Human Resource 2012 #71).

Dano, A. M. (2005). Road Injuries and Long-Run Effects on Income and Employment. *Health Economics 14*(9), 955–970.

Davis-Friedmann, D. (1991). *Long Lives: Chinese Elderly and the Communist Revolution*. Stanford University Press.

DeCicca, P. (2010). Health Insurance Availability and Entrepreneurship. *SSRN Working Paper Series*.

Donald, S. G. and K. Lang (2007). Inference with Difference-In-Differences and Other Panel Data. *The Review of Economics and Statistics 89*(2), 221–233.

Dong, K. (2009). Medical Insurance System Evolution in China. *China Economic Review 20*(4), 591–597.

Du, W. (2010). The Impact of Financing of SMEs for Property Law System Innovation. *Journal of Henan Administrative Institute of Politics and Law 4*, 023.

Earle, J. S. and Z. Sakova (2000). Business Start-Ups or Disguised Unemployment? Evidence on the Character of Self-Employment from Transition Economies. *Labour Economics 7*(5), 575–601.

Evans, D. S. and B. Jovanovic (1989). An Estimated Model of Entrepreneurial Choice under Liquidity Constraints. *The Journal of Political Economy*, 808–827.

Ezzati, M., S. Vander Hoorn, C. M. Lawes, R. Leach, W. P. T. James, A. D. Lopez, A. Rodgers, and C. J. Murray (2005). Rethinking The "Diseases of Affluence" Paradigm: Global Patterns of Nutritional Risks in Relation to Economic Development. *PLoS Med 2*(5), e133.

Fairlie, R. W., K. Kapur, and S. Gates (2011). Is Employer-Based Health Insurance a Barrier to Entrepreneurship? *Journal of Health Economics 30*(1), 146–162.

Fan, G., X. Wang, and H. Zhu (2003). NERI Index of Marketization of China's Provinces. *National Economic Research Institute, Beijing*.

Fan, G., X. Wang, and H. Zhu (2011). The Marketization Index of China: The Process of Regional Marketization Report. *Beijing: Press of Economic Science [in Chinese]*.

Feng, X. L., M. Pang, and J. Beard (2014). Health System Strengthening and Hypertension Awareness, Treatment and Control: Data from the China Health and Retirement Longitudinal Study. *Bulletin of the World Health Organization 92*(1), 29–41.

Giles, J., D. Wang, and W. Cai (2011). The Labor Supply and Retirement Behavior of China's Older Workers and Elderly in Comparative Perspective. *World Bank Policy Research Working Paper Series*.

Gruber, J. and B. C. Madrian (2002). Health Insurance, Labor Supply, and Job Mobility: A Critical Review of the Literature. *NBER Working Paper Series*, 8817.

Halla, M. and M. Zweimüller (2013). The Effect of Health on Earnings: Quasi-Experimental Evidence from Commuting Accidents. *Labour Economics 24*, 23–38.

Hamersma, S. and M. Kim (2009). The Effect of Parental Medicaid Expansions on Job Mobility. *Journal of Health Economics 28*(4), 761–770.

Hamilton, B. H. (2000). Does Entrepreneurship Pay? An Empirical Analysis of the Returns of Self-Employment. *Journal of Political Economy 108*(3), 604.

Han, S. L. (2007). Secured Transactions Law Reform in China: Can a Commercial Law Serve the Needs of the Market? *Secured Lender 63*(4), 30.

Heim, B. T. and I. Z. Lurie (2010). The Effect of Self-Employed Health Insurance Subsidies on Self-Employment. *Journal of Public Economics 94*(11), 995–1007.

Hirano, K. and G. W. Imbens (2001). Estimation of Causal Effects Using Propensity Score Weighting: An Application to Data on Right Heart Catheterization. *Health Services and Outcomes Research Methodology 2*(3-4), 259–278.

Hirano, K., G. W. Imbens, and G. Ridder (2003). Efficient Estimation of Average Treatment Effects Using the Estimated Propensity Score. *Econometrica 71*(4), 1161–1189.

Holtz-Eakin, D., J. R. Penrod, and H. S. Rosen (1996). Health Insurance and the Supply of Entrepreneurs. *Journal of Public Economics 62*(1), 209–235.

Hou, Z., Q. Meng, and Y. Zhang (2016). Hypertension Prevalence, Awareness, Treatment, and Control Following China's Healthcare Reform. *American Journal of Hypertension 29*(4), 428–431.

Imbens, G. (2014). Instrumental Variables: An Econometrician's Perspective. *NBER Working Paper*.

Izzo, J. L., D. Levy, and H. R. Black (2000). Importance of Systolic Blood Pressure in Older Americans. *Hypertension 35*(5), 1021–1024.

Jiang, Y. and F. Yi (2014). Scenario Simulation of China's Rural Workers Covered by the Urban Medical Insurance: Generational Accounting Perspective. *Economic Science* (5), 79–89.

Kannel, W. B., T. Gordon, and M. J. Schwartz (1971). Systolic versus Diastolic Blood Pressure and Risk of Coronary Heart Disease: The Framingham Study. *The American Journal of Cardiology 27*(4), 335–346.

Kapur, K. (1998). The Impact of Health on Job Mobility: A Measure of Job Lock. *Industrial & Labor Relations Review 51*(2), 282–298.

Kessler, R. C., C. Barber, H. G. Birnbaum, R. G. Frank, P. E. Greenberg, R. M. Rose, G. E. Simon, and P. Wang (1999). Depression in the Workplace: Effects on Short-Term Disability. *Health Affairs 18*(5), 163–171.

Kihlstrom, R. E. and J.-J. Laffont (1979). A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion. *The Journal of Political Economy*, 719–748.

La Porta, R. and A. Shleifer (2008). The Unofficial Economy and Economic Development. *Brookings Papers on Economic Activity* (2).

Le, A. T. (1999). Empirical Studies of Self-Employment. *Journal of Economic Surveys 13*(4), 381–416.

Levine, R. and Y. Rubinstein (2015). Smart and Illicit: Who Becomes an Entrepreneur and Does It Pay. *NBER Working Paper Series* (19276).

Liang, L. and J. C. Langenbrunner (2013). The Long March to Universal Coverage: Lessons from China. *UNICO Studies Series No.9*.

Lin, W., G. G. Liu, and G. Chen (2009). The Urban Resident Basic Medical Insurance: A Landmark Reform towards Universal Coverage in China. *Health Economics 18*(S2), S83–S96.

Liu, G. (2008). Change in Collateral System: From 1995 Collateral Law to 2007 Property Law [in Chinese]. *The Pacific Journal*.

Liu, H. and Z. Zhao (2014). Does Health Insurance Matter? Evidence from China's Urban Resident Basic Medical Insurance. *Journal of Comparative Economics*.

Liu, X. and C. Zhu (2014). Will Knowing Diabetes Affect Labor Income? Evidence from a Natural Experiment. *Economics Letters 124*(1), 74–78.

Liu, Y. (2002). Reforming China's Urban Health Insurance System. *Health Policy 60*(2), 133–150.

Madrian, B. (1994). Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job-Lock? *The Quarterly Journal of Economics*.

Marechal, V., P. Tekin, and H. Guliyeva (2009). China's New Property Rights Law – an Important Step Towards Improving Access to Credit for Small and Medium Enterprises. *The World Bank Doing Business Case Study: China*.

Ministry of Health (2013). *China Health Statistics 2013 Yearbook*. The Central People's Government of the People's Republic of China.

Mohapatra, S., S. Rozelle, and R. Goodhue (2007). The Rise of Self-Employment in Rural China: Development or Distress? *World Development*.

Norton, E. C., H. Wang, and C. Ai (2004). Computing Interaction Effects and Standard Errors in Logit and Probit Models. *Stata Journal 4*, 154–167.

Pan, J., X. Lei, and G. G. Liu (2015). Health Insurance and Health Status: Exploring the Causal Effect from a Policy Intervention. *Health Economics*.

Parker, S. C. (2004). *The Economics of Self-employment and Entrepreneurship*. Cambridge University Press.

Peng, L., C. D. Meyerhoefer, and S. H. Zuvekas (2015). The Short-Term Effect of Depressive Symptoms on Labor Market Outcomes. *Health Economics*.

Pohl, V., C. Neilson, and F. Parro (2014). Health Shocks, Education, and Labor Market Outcomes.

Rawski, T. G. (2001). What is Happening to China's GDP Statistics? *China Economic Review 12*(4), 347–354.

Rosenbaum, P. R. and D. B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika 70*(1), 41–55.

Rossin-Slater, M., C. J. Ruhm, and J. Waldfogel (2013). The Effects of California's Paid Family Leave Program on Mothers' Leave – Taking and Subsequent Labor Market Outcomes. *Journal of Policy Analysis and Management 32*(2), 224–245.

Sarki, A. M., C. U. Nduka, S. Stranges, N.-B. Kandala, and O. A. Uthman (2015). Prevalence of Hypertension in Low- And Middle-Income Countries: A Systematic Review and Meta-Analysis. *Medicine 94*(50), e1959.

Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry into Profits,Capital, Credit, Interest, and the Business Cycle*, Volume 55. Transaction Publishers.

Seuring, T., O. Archangelidi, and M. Suhrcke (2015). The Economic Costs of Type 2 Diabetes: A Global Systematic Review. *Pharmacoeconomics 33*(8), 811–831.

Stamler, J. (1998). Setting the TONE for Ending the Hypertension Epidemic. *JAMA 279*(11), 878–879.

State Council (2012). *Major Work Arrangements for Deepening the Health Care System Reform in 2012*. The Central People's Government of the People's Republic of China.

State Council (2016). *Guideline on the Integration of Basic Medical Insurance for Urban Employees and the New Rural Cooperative Medical Scheme*. The Central People's Government of the People's Republic of China.

Tang, W. (2015). Design and Effects of Rural Chronic Care Delivery Integration System from Perspective of Complex System Theory: Evidence from a Multi-faceted Intervention in Qianjiang [in Chinese]. *Huazhong University of Science and Technology Doctoral Dissertation*.

Taylor, M. P. (2001). Self-Employment and Windfall Gains in Britain: Evidence from Panel Data. *Economica*, 539–565.

The China National Bureau of Statistics (2012). *China Township Statistics [Zhongguo Jianzhixian Tongji Ziliao]*. China Statistics Press.

The Chongqing Bureau of Statistics (2005). *Chongqing 1 Percent Population Sample Survey [Chongqing Shi Baifenzhiyi Renkou Chouyang Diaocha Ziliao]*. China Statistics Press.

Tokman, V. E. (1992). *Beyond Regulation: The Informal Economy in Latin America*.

Trialists' Collaboration, F. Turnbull, B. Neal, T. Ninomiya, C. Algert, H. Arima, F. Barzi, C. Bulpitt, J. Chalmers, R. Fagard, et al. (2008). Effects of Different Regimens to Lower Blood Pressure on Major Cardiovascular Events in Older and Younger Adults: Meta-Analysis of Randomised Trials. *The BMJ 336*(7653), 1121–3.

Trialists' Collaboration, F. Turnbull, B. Neal, T. Ninomiya, C. Algert, H. Arima, F. Barzi, C. Bulpitt, J. Chalmers, R. Fagard, et al. (2014). Blood Pressure-Lowering Treatment Based on Cardiovascular Risk: A Meta-Analysis of Individual Patient Data. *The Lancet 384*(9943), 591–598.

Van Praag, C. M. and J. S. Cramer (2001). The Roots of Entrepreneurship and Labour Demand: Individual Ability and Low Risk Aversion. *Economica 68*(269), 45–62.

Velamuri, M. (2012). Taxes, Health Insurance, and Women's Self-Employment. *Contemporary Economic Policy 30*(2), 162–177.

Wagstaff, A. (2007). The Economic Consequences of Health Shocks: Evidence from Vietnam. *Journal of Health Economics 26*(1), 82–100.

Wang, D. (2009). Development of Commerical Health Insurance in China under the New Medical Reform. *Insurance Studies*.

Wang, J. (2014). The Household Level Impact of Public Health Insurance: Evidence from the Urban Resident Basic Medical Insurance in China. *Working Paper*.

Weil, D. N. (2007). Accounting for the Effect of Health on Economic Growth. *The Quarterly Journal of Economics*, 1265–1306.

Wellington, A. J. (2001). Health Insurance Coverage and Entrepreneurship. *Contemporary Economic Policy 19*(4), 465–478.

Williams, B., L. H. Lindholm, and P. Sever (2008). Systolic Pressure is All that Matters. *The Lancet 371*(9631), 2219–2221.

World Bank (2007). Reforming Collateral Laws and Registries International Best Practices and the Case of China. *FIAS/IFC PEP Report*.

World Health Organization (2009). *Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks.* The World Health Organization.

Wu, S. (2003). The Effects of Health Events on the Economic Status of Married Couples. *Journal of Human Resources 38*(1), 219–230.

Wu, Y., R. Huxley, L. Li, V. Anna, G. Xie, C. Yao, M. Woodward, X. Li, J. Chalmers, R. Gao, et al. (2008). Prevalence, Awareness, Treatment, and Control of Hypertension in China Data from the China National Nutrition and Health Survey 2002. *Circulation 118*(25), 2679–2686.

Xu, L., Y. Wang, C. D. Collins, and S. Tang (2007). Urban Health Insurance Reform and Coverage in China Using Data from National Health Services Surveys in 1998 and 2003. *BMC Health Services Research 7*(1), 37.

Yip, W. C.-M., W. C. Hsiao, W. Chen, S. Hu, J. Ma, and A. Maynard (2012). Early Appraisal of China's Huge and Complex Health-Care Reforms. *The Lancet 379*(9818), 833–842.

Zhang, B., F. Zhai, S. Du, and B. Popkin (2014). The China Health and Nutrition Survey, 1989–2011. *Obesity Reviews 15*(S1), 2–7.

Zhang, Y., W. Tang, Y. Zhang, L. Liu, and L. Zhang (2017). Effects of Integrated Chronic Care Models on Hypertension Outcomes and Spending: A Multi-town Clustered Randomized Trial in China. *BMC Public Health*.

Zhang, Y. and A. Van Stel (2016). Who should be Running Ahead? The Roles of Two Types of Entrepreneursh in China's Contemporary Economy. *Working Paper*.

APPENDIX A

CLUSTER V.S. BOOTSTRAP

Table A1: Robustness Checks for Standard Errors

	(1)	(2)	(3)	(4)
	Robust	Cluster at Town	Cluster at Village	Bootstrap
Change in blood pressure	-5.209**	-5.209***	-5.209*	-5.209**
	(2.471)	(1.560)	(2.700)	(2.355)
Age	0.0292	0.0292	0.0292	0.0292
	(0.0428)	(0.0402)	(0.0498)	(0.0456)
Female ratio	0.00168	0.00168	0.00168	0.00168
	(0.00939)	(0.00872)	(0.00641)	(0.00923)
Log(income-12)	-0.801	-0.801	-0.801	-0.801
	(0.692)	(0.835)	(0.776)	(0.547)
Baseline Blood Pressure	-0.378**	-0.378***	-0.378**	-0.378**
	(0.160)	(0.134)	(0.181)	(0.153)
Observations	541	541	541	541

Notes: Outcome variable is the percentage change in annual income (%). All regressions include basic individual characteristics. The specification used is include health variables and exclude town variables. Standard errors are not clustered in column (1), clustered at town level in column (2), clustered at village level in column (3), and bootstrapped in column (4). Significance level: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX B

USING TWO INSTRUMENTS

This appendix section presents supplementary result using two instruments $IV1$ and $IV2$: $IV1 = 1$ if the individuals are treated by intervention 1, and $IV2 = 1$ if the individuals are treated by intervention 2. Specifically, we perform the following empirical analysis:

$$\Delta bp_i = \gamma_1 + \gamma_2 IV1_i + \gamma_3 IV2_i + \gamma_4 X_i + \epsilon_i \tag{B.1}$$

$$\Delta LogInc_i = \rho_1 + \rho_2 \Delta bp_i + \rho_3 X_i + \upsilon_i \tag{B.2}$$

where X is a vector of variables including Cov, $logInc12$, and $BP12$. Robust standard errors are clustered at village level.

Main results and placebo test results are reported in Table B1 and Table B2, which are almost identical as the results using only one instrument.

Table B1: IV Regressions Using Two Instruments

	(2) IV	(3) IV	(4) IV	(5) IV
Sample: Not Disabled				
Change in blood pressure	-4.463***	-5.069***	-4.292***	-4.831***
	(1.717)	(1.962)	(1.616)	(1.817)
Age	-0.00127	-0.00477	0.0188	0.0169
	(0.0462)	(0.0473)	(0.0472)	(0.0482)
Female ratio	0.00121	0.000701	0.00139	0.000738
	(0.00611)	(0.00644)	(0.00623)	(0.00656)
Log(income-12)	-0.480	-0.384	-0.605	-0.526
	(0.693)	(0.662)	(0.745)	(0.721)
Baseline Blood Pressure	-0.343***	-0.380***	-0.331***	-0.363***
	(0.107)	(0.125)	(0.100)	(0.114)
Observations	735	735	708	708
Include Health Variables	No	No	Yes	Yes
Include Town Variables	No	Yes	No	Yes
Sample: Farming				
Change in blood pressure	-6.557**	-8.825**	-6.394**	-8.841**
	(3.030)	(4.116)	(2.884)	(3.786)
Age	0.0123	-0.0153	0.0110	-0.0164
	(0.0626)	(0.0707)	(0.0623)	(0.0753)
Female ratio	0.00437	0.00455	0.00212	0.00126
	(0.0118)	(0.0134)	(0.0120)	(0.0139)
Log(income-12)	-0.306	-0.0888	-0.530	-0.316
	(0.687)	(0.704)	(0.751)	(0.782)
Baseline Blood Pressure	-0.462**	-0.621**	-0.449**	-0.615**
	(0.217)	(0.285)	(0.217)	(0.267)
Observations	566	566	541	541
Include Health Variables	No	No	Yes	Yes
Include Town Variables	No	Yes	No	Yes

Notes: Outcome variable is the percentage change in annual income (%). All regressions include basic individual characteristics and use PS weighting. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. "Farming" refers to individuals who are engaged in agriculture work. "Not Disabled" refers to individuals whose physical ability are not severely limited due to diseases. Robust standard errors are clustered at village level. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table B2: Placebo Tests Using Two Instruments

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV
Sample: Disabled					
Change in blood pressure	-0.534	-2.521	-0.296	-2.575	-0.875
	(0.548)	(2.198)	(2.138)	(2.094)	(2.069)
Age	0.0778	0.0992**	0.0897*	0.0885*	0.0799*
	(0.0475)	(0.0489)	(0.0463)	(0.0465)	(0.0459)
Female ratio	0.0126	0.0148	0.0148	0.0117	0.0125
	(0.0115)	(0.0111)	(0.0111)	(0.0111)	(0.0112)
Log(income-12)	-0.424	0.0686	-0.438	0.0231	-0.422
	(0.731)	(0.782)	(0.709)	(0.737)	(0.697)
Baseline Blood Pressure	-0.157***	-0.283**	-0.140	-0.287**	-0.176
	(0.0417)	(0.129)	(0.130)	(0.124)	(0.126)
Observations	331	342	331	342	331
Include Health Variable	Yes	No	Yes	No	Yes
PS Weighting	No	No	No	Yes	Yes
Sample: Not Farming					
Change in blood pressure	-0.404	0.0970	1.199	0.0969	1.062
	(0.428)	(1.336)	(1.901)	(1.236)	(1.749)
Age	0.0552	0.0439	0.0527	0.0511	0.0547
	(0.0400)	(0.0471)	(0.0466)	(0.0380)	(0.0373)
Female ratio	0.0114	0.0104	0.0127	0.0101	0.0129
	(0.0107)	(0.00928)	(0.0108)	(0.00970)	(0.0113)
Log(income-12)	0.109	0.357	0.0767	0.346	0.106
	(0.502)	(0.579)	(0.460)	(0.603)	(0.487)
Baseline Blood Pressure	-0.117***	-0.107	-0.0308	-0.109	-0.0403
	(0.0404)	(0.0778)	(0.114)	(0.0698)	(0.102)
Observations	498	511	498	511	498
Include Health Variable	Yes	No	Yes	No	Yes
PS Weighting	No	No	No	Yes	Yes

Notes: Outcome variable is the percentage change in annual income (%). All regressions include basic individual characteristics and use PS weighting. "Health variables" include salt control, fat control, anti-hypertension drug usage, and inpatient records. "Town variables" include the number of residents in the town and the distance from the town to county seat. "Farming" refers to individuals who are engaged in agriculture work. "Not Disabled" refers to individuals whose physical ability are not severely limited due to diseases. Robust standard errors are clustered at village level. Significance level: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX C

DATA DISCUSSION

We use five waves of the China Health and Nutrition Survey (CHNS) data: 2000, 2004, 2006, 2009, and 2011. It is a panel dataset and individuals were interviewed in each wave. Due to attrition, new individuals were added to maintain the representativeness of the sample. 91.37% of the observations appeared in more than one wave, and 38.57% appeared in all five waves, as shown in Figure C1. The annual attrition rate is around 23%, which is high relative to U.S. datasets. I show in Table C1 the comparative statistics for individual characteristics between consecutive waves. The value reported is the mean value for the baseline year, and the significance level is for the t-test comparison between the baseline year and the next consecutive wave. For example, column (1) reports the mean values for observations in the 2000 wave, and the comparison is between 2000 wave and 2004 wave. Besides 2000 wave, the other waves all look very similar, thus the representativeness of the sample is generally maintained.

Figure C1: Frequency of Appearance in Five CHNS Waves

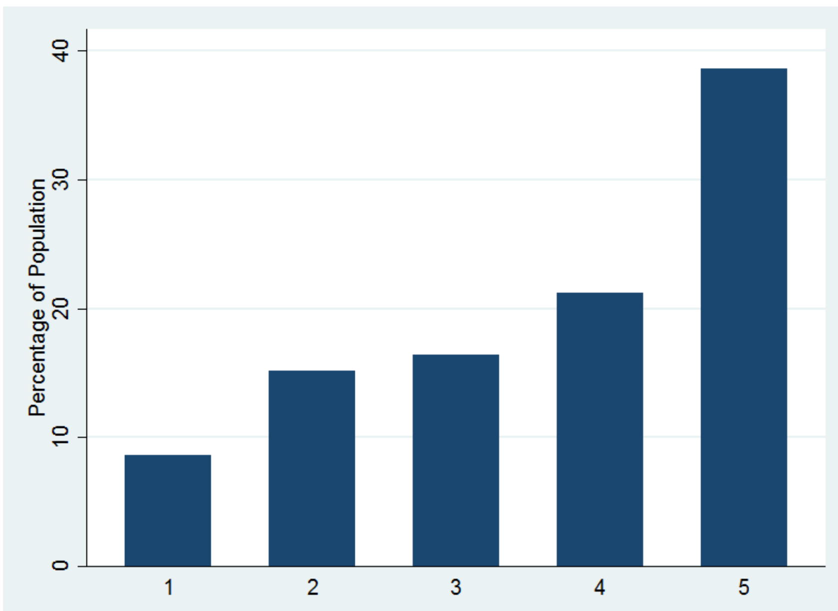


Table C1: Comparative Statistics for Different CHNS Waves

	(1) 2000 vs 2004	(2) 2004 vs 2006	(3) 2006 vs 2009	(4) 2009 vs 2011
Age (year)	39.45**	40.62*	41.62	41.97
Female Share (%)	46.93	45.10	44.14	42.63
Education (year)	8.72***	9.79	9.98	10.14
Marriage Rate (%)	79.78**	83.74	85.48	84.16
Poor Health (%)	11.14***	21.26***	16.40	16.88
Log(Income)	7.65	7.58***	8.07***	8.88
Spouse Works (%)	66.06***	59.32	60.14*	56.13
Urban Hukou	0.60***	0.68*	0.64	0.65
Senior Professional (%)	6.47	7.78	8.17	8.28
Junior Professional (%)	6.69	6.40	6.00	6.52
Administrator (%)	7.23*	9.56	8.50	7.56
Office Staff (%)	8.42	8.51	9.95	10.76
Skilled Worker (%)	11.92	11.74	9.82	10.56
Non-skilled Worker (%)	9.62	11.08	12.46	11.41
Police or Army (%)	1.11	0.79	0.33	0.46
Driver (%)	2.97*	4.35	4.61	3.91
Service Worker (%)	11.30	12.40*	15.29	17.47**

Notes: The value reported is the mean value for the baseline year, and the significance level is for the t-test comparison between the baseline year and the next consecutive wave. For example, column (1) reports the mean values for observations in the 2000 wave, and the comparison is between the 2000 wave and the 2004 wave. Significance level: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

APPENDIX D

ALTERNATIVE EMPIRICAL STRATEGY

In order to rule out the possibility that the relative increase in self-employment rate for urban Hukou is driven by other events, instead of the implementation of URBMI, we present here an alternative empirical strategy. We focus on the transition from a salaried position to self-employment among urban Hukou residents, separately during the periods before and after the implementation of URBMI. We use a difference-in-differences model to identify that job lock existed before 2009, and then another one after 2009 to show that job lock disappeared after the implementation of URBMI.

We observe in Table 3.1 that before the implementation of URBMI, the transition rate for urban Hukou residents is much higher for individuals with employer-provided health insurance compared to those without. In this section, we estimate the entrepreneurship barrier using a difference-in-differences strategy following Holtz-Eakin et al. (1996) and Fairlie et al. (2011): comparing the effect of employment-based insurance on self-employment transition rate between those with low and high demand for the employment-based health insurance.

The most commonly used indicator in the literature to measure the demand for employer-provided health insurance is whether employees' spouses have insurance coverage that can cover workers.[1] However, China's employer-provided insurance program, UEBMI, does not cover employees' spouse and children, thus spouse's insurance coverage can not be used

[1] Other indicators used in the literature, like having a pregnant wife, having family member in poor health, and having a large family, usually failed to generate significant results. We used these indicators as well. The point estimations are as predicted, but not significant. Thus we will not report these results here.

to measure the demand for employment-based health insurance in China. Instead, we use whether an individual has non-employment based insurance as the indicator. If a wage-worker has non-employment based insurance plans, he/she may have lower demand for employer-provided health insurance, compared to those who do not have non-employment based insurance. The effect of employer-based insurance on transitions to self-employment can be confounded by job quality, but by taking the difference of the differences between individuals with and without non-employment based insurance, we can mitigate this endogeneity. Among individuals with employer-provided health insurance in our sample, 17.6% had supplemental non-employment based insurance. Among individuals without employer-provided health insurance, 31.6% had non-employment based insurance. In China, the most common non-employment based health insurance before the implementation of URBMI was commercial insurance. In our sample, 72.7% of individuals with non-employment based insurance in the pre-policy period had commercial insurance.

Empirically, we estimate the following model for pre-policy ($t < 2009$) and post-policy ($t \geq 2009$) respectively:

$$
\begin{aligned}
1[NewSelf]_{it} = \beta_0^{pre} + \beta_1^{pre} 1[EI]_{it} + \beta_2^{pre} 1[HD]_{it} + \beta_3^{pre}(1[EI]_{it} \cdot 1[HD]_{it}) + \\
\beta_4^{pre} X_{it} + \gamma_t^{pre} + \tau_c^{pre} + \epsilon_{it}^{pre}, \;\; t < 2009 \\
1[NewSelf]_{it} = \beta_0^{post} + \beta_1^{post} 1[EI]_{it} + \beta_2^{post} 1[HD]_{it} + \beta_3^{post}(1[EI]_{it} \cdot 1[HD]_{it}) + \\
\beta_4^{post} X_{it} + \gamma_t^{post} + \tau_c^{post} + \epsilon_{it}^{post}, \;\; t \geq 2009
\end{aligned}
\tag{D.1}
$$

where $1[NewSelf]_{it} = 1$ indicates a transition from a wage-earner in wave t to an entrepreneur in wave $t+1$, and $1[EI]_{it} = 1$ indicates individual i have employment-based health insurance in wave t. $1[HD]_{it} = 1$ indicates having a *high* demand for employer-provided health insurance in wave t, defined as *not* having supplementary non-employment-based insurance in wave t. The interaction term $(1[EI]_{it} \cdot 1[HD]_{it})$ provides the difference-in-differences estimator. The other aspects of the model are similar to equation 3.1. Negative coefficients of the interaction term (β_3^{pre} and β_3^{post}) are consistent with the notion of entrepreneurship lock.

In an ideal world where non-employment based health insurance is randomly assigned, this difference-in-differences strategy can give us an accurate estimation of entrepreneurship lock. However, in our case, the assignment of individuals with and without non-employment based insurance plan is not a random mechanism. Individuals with and without non-employment based insurance can differ in health status, risk preference, and wealth, all of which can affect their labor market decisions. We conducted t-test and presented the comparative statistics of these variables in the pre-policy period in Table D1 Panel A. In general, there are no significant differences in key characteristics between individuals with and without non-employment based health insurance, except for the indicator for whether a spouse has a job, the dummy for being a senior professional, and the dummy for being a service worker. Nevertheless, we discuss below three potential concerns of violating the random assignment assumption and how it may affect our estimation strategy.

First, individuals can differ in health condition, and it is possible that those who purchased non-employment based health insurance have worse health than those who did not purchase additional insurance. We compared the health indicators discussed above between those with and without non-employment based insurance and did not find statistically significant differences. We acknowledge that there may be unobserved differences in health conditions. However, if workers with non-employment based health insurance had worse health status than those without, we would underestimate the effect of employment-based insurance on job lock. In particular, individuals with poor health are less mobile, thus less likely to transition into self-employment. But those with supplementary non-employment based insurance have low demand for employment-based insurance and are therefore not as subject to job lock as those without supplementary health insurance. That is, the potential selection, if any, works opposite direction as the main effect, and therefore it could underestimate the effect.

Second, individuals with more insurance might be more risk-averse than those with less insurance. We use the usage of cigarette and alcohol as an indicator of risk preference. Heavy drinkers and heavy smokers are usually more reckless, thus less risk averse. We categorize risk aversion by identifying whether an individual smokes more than 20 cigarettes per day, or drink more than 10 bottles of beer or 1 liter of liquor each week. We conducted a t-test to compare this risk aversion indicator and did not find statistically significant difference between those with and without non-employment based insurance (30.90% vs 26.66%; p-value of 0.12). Even if risk preference matters, workers with alternative non-employment based insurance are more risk averse and should value their employer-provided insurance more. This is again the opposite direction as our theory proposed, so our estimation can be considered as a conservative estimation.

Third, individuals with non-employment insurance may be richer than those without this insurance. It is unclear how wealth affects transition rates to self-employment. On one hand, it's easier for richer people to break the barrier to become entrepreneurs because they have more money. On the other hand, a high wage also means a high opportunity cost to leave the current jobs. Nevertheless, there is no statistically significant difference in the logarithm of annual wage and household income between those with non-employment based insurance and those without such insurance.

In sum, we do not observe significant differences in main characteristics between individuals with and without non-employment based health insurance, except for three indicators, one for having working spouses and the other two for being senior professionals or service workers. To further balance the two groups, we also utilize the propensity score method with an inverse probability of treatment weighting (PS Weights). The weighted characteristics for each group are reported in Panel B of Table D1. After using PS weights, all key characteristics are balanced.

We report estimation results with and without PS weights in Table D2, with pre-policy regressions in column (1)-(2) and post-policy regressions in column (3)-(4). Before URBMI was implemented, for the overall population, having a high demand for employer-provided health insurance increases the difference in the rates of transition to self-employment between individuals with and without employer-provided health insurance by 7.65 percentage points, or 151% compared to the baseline transition rate. For male workers, the effect is robust but larger. In the post-policy period, not only do some coefficients become positive, but the standard errors are also much larger. This finding suggests that there is no evidence of entrepreneurship lock after the implementation of URBMI. The existence of entrepreneurship lock in the pre-policy period and its disappearance in the post-policy period also imply that the non-employment based health insurance alleviated the job lock, and succeeded in establishing individual coverage as a legitimate alternative to employer-sponsored coverage.

Table D1: Comparative Statistics for Individuals with and Without Non-Employment Insurance

	(1) With NEBI	(2) Without NEBI	(3) Difference	(4) Count
Panel A: Without PS Weights				
Age (year)	40.48	39.72	0.76	1488
Female Share (%)	43.15	40.85	2.30	1491
Education (year)	11.07	11.30	-0.24	1378
Marriage Rate (%)	85.76	84.91	0.84	1477
Poor Health (%)	14.29	14.98	-0.70	1491
Spouse Has Job (%)	66.47	60.98	5.50*	1491
Log(Annual Income)	8.94	8.81	0.13	1491
Share as Senior Professional (%)	18.13	14.04	4.09*	1489
Share as Service Worker (%)	5.85	10.99	-5.14**	1489
Panel B: With PS Weights				
Age (year)	40.39	40.21	0.18	1488
Female Share (%)	40.76	41.29	-0.53	1491
Education (year)	11.15	11.23	-0.07	1378
Marriage Rate (%)	84.88	85.16	-0.28	1477
Poor Health (%)	16.54	15.90	0.64	1491
Spouse Has Job (%)	61.24	62.47	-1.23	1491
Log(Annual Income)	8.88	8.59	0.29	1491
Share as Senior Professional (%)	14.09	14.72	-0.63	1489
Share as Service Worker (%)	10.98	10.40	0.58	1489

Notes: (1) Column (1) are for individuals with non-employment based insurance (with NEBI), which represents low demand for employer-provided insurance. Column (2) represents high demand. (2) We only report the t-test results for two of the seven major occupational dummies in this table, since the results for the others are all insignificant even without PS weights. (3) Panel A reports results without the inverse probability of treatment weighting. Panel B reports results with it. (4) Significance level in t-test: *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table D2: Alternative Strategy: Existence of Entrepreneurship Lock

	(1)	(2)	(3)	(4)
	Pre-policy		Post-policy	
HaveEmpIns*HighDemand with PS Weights	-0.0765*	-0.113**	-0.0343	-0.0871
	(0.0424)	(0.0474)	(0.0924)	(0.115)
HaveEmpIns*HighDemand	-0.0665**	-0.0953**	0.0712	0.0406
	(0.0338)	(0.0427)	(0.0622)	(0.123)
Observations	1,334	786	498	303
Dependent Mean	0.0507	0.0582	0.0456	0.0493
Sample	All	Male	All	Male

Notes: (1) This table reports the treatment effects of the interaction term (HaveEmployerIns*HighDemand, i.e. having employer-provided health insurance and having high demand for this insurance), with and without inverse probability of treatment weights (PS Weights). The dependent variable $1[NewSE] = 1$ indicates transition from a wage-earner in current wave to an entrepreneur in the next wave. (2) All specifications include year and city fixed effects. (3) Significance level: *** p<0.01, ** p<0.05, * p<0.1

APPENDIX E

CHARACTERISTICS OF THE TWO TYPES OF SELF-EMPLOYMENT

I analyze the characteristics of the two types of self-employment–own-account workers and owner-managers. I examine their income profile and address the selection in the occupational choice. I separate the population into three categories: wage earners, self-employed own-account workers, and self-employed owner-managers. I found that: (1) Under a multinomial logistic model, own-account workers and self-employed owner-managers show vastly different characteristics. The former are often not well educated, not from well-educated families, and with low income. On the opposite, the owner-managers are richer and come from more educated families. (2) Using a similar method as Hamilton (2000), I found that own-account workers earned lower salary when they were wage workers compared to the remaining wage earners. After their transition into self-employment, their income decreased even further. For owner-mangers who were salary workers in the past, there is no difference in their wage compared to the remaining wage earners. After the transition, their income increased significantly. Both findings are evidence that self-employed own-account workers' entrepreneurship reflects some distress or survival necessity, while self-employed owner-manger's entrepreneurship is an unconstrained choice.

E.1 A MULTINOMIAL LOGISTIC MODEL

I utilize a multinomial logit model, and particularly focus on characteristics that tent to associate with either the "necessity entrepreneur view" or the "opportunity entrepreneur view"– education, family background, and household income. I further conduct pooling tests for the differences across the categories, which suggests that the difference within self-employment sectors is even bigger than the difference between sectors.

The multinomial logit model uses salary workers as the baseline category.

$$log\frac{P_{ijt}}{P_{iWt}} = \alpha_j + \beta_j X_{it} + \gamma_{1j} City_{it} + \gamma_{2j} Year_{it} + \epsilon_{ijt} \tag{E.1}$$

P_{ijt} stands for the probability that person i is own-account self-employed worker ($j = 1$), or self-employed owner manager ($j = 2$) at time t, and P_{iWt} denotes the probability that the person is a wage worker in time t. X_{it} is a vector of regressors that includes demographics (age, age square, gender, marriage status), years of schooling, the average number of hours worked per year, whether have health insurance, occupational status, log of household income, log of individual income, and parents' education status. City and year fixed effects are included.

I also estimated similar model using characteristics of the previous wave since there are concerns that current occupational status is determined by previous characteristics (denoted in Table E2 as with time lag):

$$log\frac{P_{ijt}}{P_{iWt}} = \alpha_j + \beta_j X_{i,t-1} + \gamma_{1j} City_{it} + \gamma_{2j} Year_{it} + \epsilon_{ijt} \tag{E.2}$$

Results for the marginal effect are reported in Table E2, showing significant differences among each group. For own-account workers, the coefficients on own education, Hukou status, household income, individual income, and health insurance status are all significantly negative. This means that compared to salary workers, they are less educated, poorer, rural Hukou and no health insurance. Thus individuals associated with disadvantaged characteristics are more likely to become own-account workers. For owner-managers, although they are more likely to be less educated and have no health insurance, they have higher household income. The effect for each covariates is also smaller. For example, when I increase one year of school, my probability of being own-account workers decreases by 0.24%, and my probability of becoming owner-manager only decreases by 0.12%. The results are robust for city residents with urban Hukou.

I further conduct Cramer-Ridder pooling test to check whether the differences between each category in the previous multinomial logit model are significant. The null hypothesis is that the two states have the same regressors coefficient apart from the intercept.[1] Results in Table E1 show that this hypothesis ($prob(P > \chi^2) = 0.000$) is rejected for every group, i.e. the three categories are very different from each other. What is phenomenal is that the biggest difference is between own-account workers and owner-managers. This means that the difference within self-employment sectors (own-account workers v.s. owner-managers) is even bigger than the difference between sectors (employee v.s. self-employment).

E.2 INCOME PROFILE AND SELECTION

Figure E1 shows the distribution of annual individual income (1989-2011 combined). We can see that the income of own-account workers is the lowest, and owner-managers earn at least the same amount as salary workers. I further plot the income distribution for each year in Figure E2. The annual income increased over the years, but the patterns are similar for every year: own-account workers earn the least, and owner-managers earn at least the same amount as salary workers.

However, the differences in income can come from differences in return to self-employment or it can come from selection of individuals into different sectors. Here, I use a similar strategy as in Hamilton (2000) to show that selection plays an important role in the income difference between sectors in urban China. I find that individuals who transit from salary workers to own-account workers have lower income compared to the remaining salary workers. Their income decreases further after the transition. These two findings indicate that own-account

[1] To test this hypothesis, the following test statistics can be used: $LR = 2(lnL - lnLr)$ where lnL is the maximum log likelihood of the original model and $lnLr$ the maximum log likelihood, if the estimates are constrained to be equal. LR asymptotically has a chi-square distribution with k degrees of freedom where k is the number of restrictions implied.

workers' productivity is low in both wage sector and self-employed sector, which further rejects that own-account workers are true entrepreneurs.

In order to emphasize selection, I compare the population who transitioned from wage earners in one wave to self-employment with individuals who remain to be employees. I define the transition rate as the proportion of wage-earners in one wave who became new own-account workers (or new owner-managers) in the next consecutive wave. This is done for each two-wave panels: 1989-1991, 1991-1993, 1993-1997, 1997-2000, 2000-2004, 2004-2006, 2006-2009, and 2009-2011. I also add another category–unemployed individuals–to help illustrate the results. Transition rate within different employment status are shown in Table E3 and summary statistics for individuals with different transition status are shown in Table E4. From Table E3, we can see that 87.99% of original wage-earners remain as wage-earners, 7.69% transition to be own-account workers and 1.74% transition to be self-employed owner-managers. Based on the mean annual income in the summary statistics, workers remaining in paid employment earn much more than those who transit to be own-account workers (7076 v.s. 4677 RMB), but lower than the would-be entrepreneurs prior to entering self-employment (7076 v.s. 8905 RMB).

To further examine the statistical significance of this relationship, I implement a post-program estimator as in Hamilton (2000). In particular, I estimate wage regressions for paid employees including a dummy variable that indicating their sector status in the next consecutive wave:

$$logwage_{it} = \alpha_0^1 + \alpha_1^1 OwnAccountNext_{it} + \beta_1 X_{it} + \gamma_1 City_{it} + \eta_1 Year_{it} + \epsilon_{it} \tag{E.3}$$

$$logwage_{it} = \alpha_0^2 + \alpha_1^2 ManagerNext_{it} + \beta_2 X_{it} + \gamma_2 City_{it} + \eta_2 Year_{it} + \epsilon_{it} \tag{E.4}$$

$$logwage_{it} = \alpha_0^3 + \alpha_1^3 UnemployNext_{it} + \beta_3 X_{it} + \gamma_2 City_{it} + \eta_3 Year_{it} + \epsilon_{it} \tag{E.5}$$

$UnemployNext_{it}$, $OwnAccountNext_{it}$ and $ManagerNext_{it}$ equals 1 if the individual transitions from a salary worker in wave t to unemployed/own-account worker/owner-manager in wave $t+1$, and equals 0 if remaining to be a salary worker. I use two measures for the dependent variable log of wage income: log of hourly wage and log of annual wage. X_{it} is a vector of individual characteristics in wave t. City and year fixed effects are included.

Results are shown in Table E5. The coefficients of $ManagerNext_{it}$ are insignificant, while the coefficients of $OwnAccountNext_{it}$ and $UnemployNext_{it}$ are significantly negative. This means that those who transit to be own-account workers or become unemployed originally earn much less salary than the remaining wage-earners. The would-be owner-managers originally had about the same amount of salary income as the workers remaining in paid employment. The patterns are consistent whether using annual wage or hourly wage as dependent variables, and whether for the overall population or for urban Hukou only. If wage can be considered as an indicator for labor productivity in the wage sector, then the new own-account workers are the less productive original employees. Notice that the wage distribution of own-account workers are very similar to the unemployed, it is very likely the new own-account workers turn to self-employment because they were forced out of the main sector.

However, this still cannot prove that these new own-account workers choose self-employment out of survival necessity, since self-employment can require different skills than being employees. A lousy employee might be a good entrepreneur. Thus, I check the changes in income after people change sectors. More specifically, I estimate the following model:

$$\Delta log(Income_{it}) = \alpha_1^j \sum Transittype_{it}^j + \beta_1 X_{it} + \gamma_1 City_{it} + \eta_1 Year_{it} + \epsilon \tag{E.6}$$

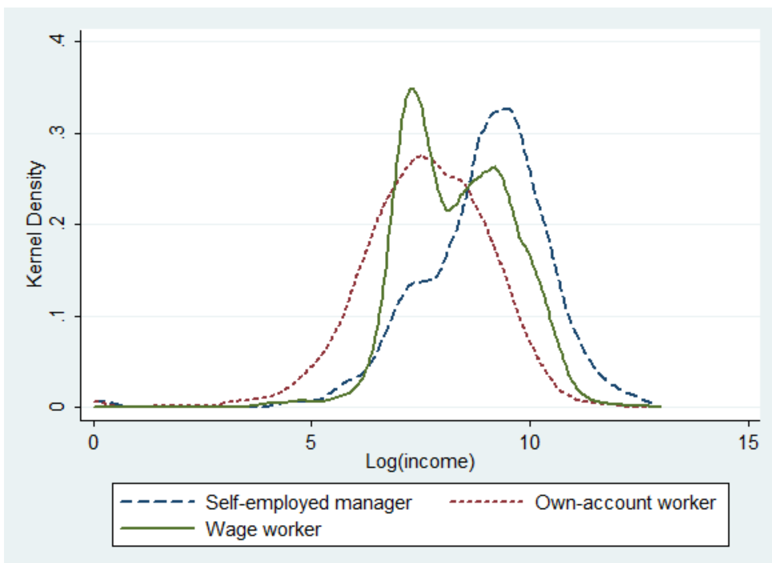
$$\Delta log(HourInc_{it}) = \alpha_2^j \sum Transittype_{it}^j + \beta_2 X_{it} + \gamma_2 City_{it} + \eta_2 Year_{it} + \epsilon \tag{E.7}$$

Change in log of annual income ($\Delta log(Income_{it})$) and change in log of hourly income ($\Delta log(HourInc_{it})$) are used as dependent variables. $Transittype_{it}^j$ is a set of dummy variable for individual i with different transition types j from wave t to wave $t+1$: from wage earner to own-account worker, from wage earner to owner-manager, from own-account worker to wage earner, from own-account worker to owner-manager, from owner-manager to wage earner, and from owner-manager to own-account worker.

Results are reported in Table E6. The first two rows reveal a clear pattern: compared to remaining wage earners, individuals' income decreased (or increased less) after changing jobs from salary workers to own-account self-employers, while individuals' income increased more after changing jobs from salary workers to self-employed owner-managers. $\Delta log(Income_{it})$ is centered around 0.303 ($sd = 1.2$), and $\Delta log(HourInc_{it})$ has a mean of 0.334 ($sd = 0.98$).

Thus magnitude wise, the differences in income changes among different transition types are not trivial.

Figure E1: Income Distribution (1989-2011 Combined)



Note: This graph shows the overall annual income distribution of self-employed owner-managers, self-employed own-account workers, and wage earners. Annual income is measured in RMB.

Figure E2: Income Distribution by Year

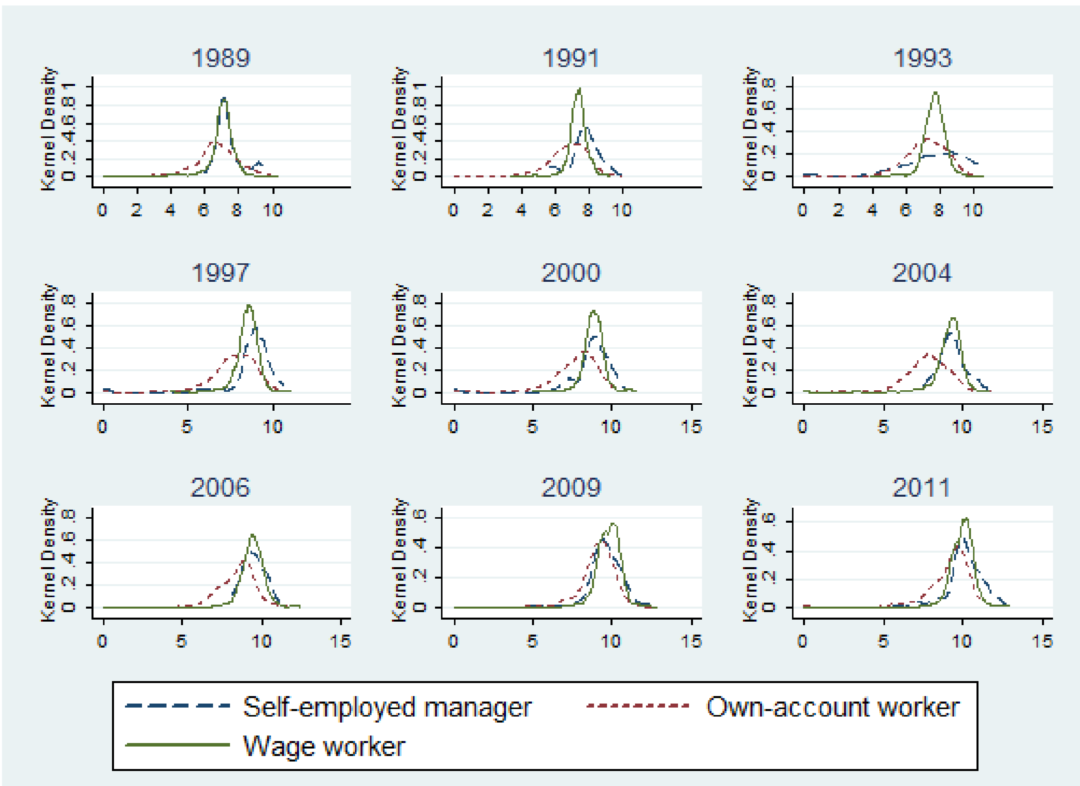


Table E1: Cramer-Ridder Test

	lnL	$lnLr$	LR	$P > \chi^2$
Own-account vs Owner-Manager	-4949	-7975	6051	0.000
Own-account vs Salary Worker	-4949	-5164	428	0.000
Owner-manager vs Salary Worker	-4949	-5170	441	0.000

Table E2: Marginal Effect of Multinomial Logit Regression

	(1) Own-account	(2) Owner-manager	(3) Own-account	(4) Owner-manager	(5) Own-account	(6) Owner-manager
Education	-0.00239***	-0.00120*	-0.00166**	-0.00128*	-0.00516***	-0.000477
	(0.000675)	(0.000630)	(0.000715)	(0.000666)	(0.00112)	(0.000684)
Age	0.00147	0.000388	0.000813	-0.000944	-0.00199	0.00198
	(0.00136)	(0.00121)	(0.00151)	(0.00123)	(0.00215)	(0.00158)
AgeSq	-1.41e-05	-9.07e-06	-7.16e-06	4.76e-06	3.74e-05	-2.78e-05
	(1.56e-05)	(1.42e-05)	(1.75e-05)	(1.42e-05)	(2.50e-05)	(1.91e-05)
Urban Hukou	-0.0277***	-0.00529			-0.0859***	-0.00393
	(0.00517)	(0.00402)			(0.00813)	(0.00480)
Female	-0.00892*	-0.00133	-0.00388	-0.000108	-0.00280	-0.00957**
	(0.00459)	(0.00351)	(0.00455)	(0.00370)	(0.00718)	(0.00435)
Insurance	-0.0365***	-0.0330***	-0.0340***	-0.0344***	-0.0705***	-0.0289***
	(0.00505)	(0.00395)	(0.00486)	(0.00426)	(0.00798)	(0.00474)
Log(Ind Income)	-0.00216	0.0107***	-0.00193	0.0105*	0.00120	0.00144
	(0.00197)	(0.00403)	(0.00212)	(0.00537)	(0.00179)	(0.00153)
Annual Work Hour	1.95e-05***	1.69e-05***	2.39e-05***	1.73e-05***	1.97e-05***	7.70e-06***
	(2.75e-06)	(2.26e-06)	(3.10e-06)	(2.79e-06)	(4.07e-06)	(2.44e-06)
Married	0.0172	-0.0106	0.00516	-0.00682	-0.0105	0.00753
	(0.0127)	(0.0101)	(0.0134)	(0.00981)	(0.0188)	(0.0120)
Parent Self-emp	0.0136	0.00753	0.0248	0.0118	-0.00993	0.00180
	(0.0134)	(0.0104)	(0.0152)	(0.0117)	(0.0199)	(0.0118)
Parent Education	-0.0178*	0.00856	-0.0279***	0.0162**	-0.0474**	0.0269***
	(0.0103)	(0.00663)	(0.0101)	(0.00734)	(0.0193)	(0.00849)
Spouse Has Job	0.0222***	0.00199	0.0133**	-0.00274	0.0203**	-0.000585
	(0.00587)	(0.00429)	(0.00543)	(0.00449)	(0.0104)	(0.00543)
Observations	11,398	11,398	8,662	8,662	7,622	7,622
Time Lag	No	No	No	No	Yes	Yes
Hukou			Urban	Urban		

Notes: Base Category is Salary Worker. Marginal effect is reported. Time and city fixed effect are added. Significance level: *** p<0.01, ** p<0.05, * p<0.1

Table E3: Transition Rate within Employment Status (%): 1989-2011 Combined

	Employment Status Next Wave				
	Unemployed	Salary Worker	OwnAccount	OwnerManager	Total
Unemployed	26.6	52.39	15.43	5.59	100
Salary Worker	2.59	87.99	7.69	1.74	100
Own-account	2.08	16.25	78.64	3.02	100
Owner-manager	5.35	39.46	28.09	27.09	100
Total Percent	4.06	62.34	31.18	2.42	100
Total Count	668	10,265	5,134	398	16,465

Table E4: Summary Statistics for Different Transition Status

	Individual Income	Annual Work Hour	Age	Education
remain salary	7076.25	2057.36	39.23	9.43
salary to ownaccount	4577.34	2078.31	38.02	6.86
salary to manager	8905.28	2133.32	38.23	9.10
remain ownaccount	3586.00	1898.60	40.44	5.30
ownaccount to salary	5139.61	1884.65	36.74	7.19
ownaccount to manager	6569.99	2354.14	36.51	7.64
remain manager	15127.43	2729.14	38.23	9.26
manager to salary	9256.94	2382.80	37.27	8.79
manager to ownaccount	14212.81	2561.08	38.07	8.12

Table E5: Self-Selection with Post-Program Estimator

	(1)	(2)	(3)	(4)	(5)	(6)
Outcome Variable: Log(Hourly Wage)						
Own-account	-0.145***			-0.130***		
	(0.0402)			(0.0488)		
Owner-manager		-0.0422			-0.0127	
		(0.0646)			(0.0728)	
Unemployed			-0.152***			-0.17***
			(0.0579)			(0.062)
Education	0.0241***	0.024***	0.0236***	0.0227***	0.0224***	0.022***
	(0.00287)	(0.0029)	(0.00287)	(0.00306)	(0.00309)	(0.0031)
Age	0.0356***	0.031***	0.0329***	0.0394***	0.0370***	0.039***
	(0.00581)	(0.0059)	(0.00582)	(0.00632)	(0.00630)	(0.0062)
Female	-0.133***	-0.14***	-0.132***	-0.126***	-0.133***	-0.13***
	(0.0165)	(0.017)	(0.0167)	(0.0173)	(0.0174)	(0.017)
Married	0.126***	0.121***	0.111***	0.123***	0.116***	0.11**
	(0.0398)	(0.040)	(0.0397)	(0.0422)	(0.0424)	(0.042)
Recent Change Job	-0.237***	-0.215***	-0.216***	-0.215***	-0.207***	-0.21***
	(0.0366)	(0.0377)	(0.0373)	(0.0392)	(0.0397)	(0.039)
Parent Education	-0.0316	-0.0389	-0.0481	-0.0158	-0.000619	-0.0055
	(0.0399)	(0.0394)	(0.0392)	(0.0417)	(0.0420)	(0.042)
Spouse Has Job	0.0432	0.0443	0.0411	0.0301	0.0423	0.034
	(0.0313)	(0.0318)	(0.0319)	(0.0324)	(0.0327)	(0.033)
Outcome Variable: Log(Annual Wage)						
Own-account	-0.104***			-0.110**		
	(0.0391)			(0.0491)		
Owner-manager		-0.0447			-0.0291	
		(0.0579)			(0.0586)	
Unemployed			-0.167**			-0.15***
			(0.0653)			(0.057)
Education	0.0253***	0.0262***	0.0256***	0.0250***	0.0248***	0.025***
	(0.00309)	(0.00312)	(0.00312)	(0.00341)	(0.00341)	(0.0034)
Age	0.0462***	0.0443***	0.0464***	0.0529***	0.0515***	0.054***
	(0.00616)	(0.00628)	(0.00639)	(0.00679)	(0.00681)	(0.0068)
Female	-0.155***	-0.162***	-0.162***	-0.150***	-0.158***	-0.16***
	(0.0166)	(0.0168)	(0.0169)	(0.0176)	(0.0177)	(0.018)
Married	0.0978**	0.0941**	0.0774**	0.0991**	0.0942**	0.086**
	(0.0387)	(0.0386)	(0.0395)	(0.0413)	(0.0413)	(0.041)
Recent Change Job	-0.200***	-0.188***	-0.183***	-0.182***	-0.179***	-0.19***
	(0.0345)	(0.0354)	(0.0354)	(0.0366)	(0.0367)	(0.036)
Parent Education	-0.0418	-0.0548	-0.0600	-0.0360	-0.0269	-0.026
	(0.0383)	(0.0382)	(0.0393)	(0.0404)	(0.0407)	(0.040)
Spouse Has Job	0.0192	0.0208	0.0194	0.00460	0.0120	0.0099
	(0.0299)	(0.0305)	(0.0307)	(0.0317)	(0.0318)	(0.032)
Observations	5,913	5,693	5,715	5,067	4,993	5,009
Sample	All	All	All	UrbanHukou	UrbanHukou	UrbanHukou

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table E6: Differences in the Change of Income

	(1) Δlog(Income)	(2) Δlog(HourlyIncome)	(3) Δlog(Income)	(4) Δlog(HourlyIncome)
wage to ownaccount	-0.248***	-0.109	-0.112	0.0523
	(0.0776)	(0.0744)	(0.114)	(0.101)
wage to manager	0.340**	0.133	0.423***	0.204*
	(0.135)	(0.0989)	(0.163)	(0.105)
ownaccount to wage	0.101	0.0738	0.131	0.0179
	(0.0870)	(0.0844)	(0.106)	(0.0951)
ownaccount to manager	0.202	-0.00957	0.184	0.149
	(0.166)	(0.164)	(0.203)	(0.183)
manager to wage	-0.251**	-0.137	-0.310**	-0.177*
	(0.0984)	(0.0871)	(0.123)	(0.103)
manager to ownaccount	-0.892***	-0.281	-0.922***	-0.433
	(0.223)	(0.202)	(0.305)	(0.311)
Education	0.00685	0.00595	0.00368	0.00656*
	(0.00422)	(0.00383)	(0.00388)	(0.00380)
Age	-0.0243***	-0.0195***	-0.0210***	-0.0126*
	(0.00745)	(0.00755)	(0.00736)	(0.00748)
Urban Hukou	0.0995**	0.0915**		
	(0.0392)	(0.0364)		
Female	-0.0342	-0.0201	-0.0184	-0.0246
	(0.0235)	(0.0215)	(0.0206)	(0.0209)
Married	-0.0497	-0.0532	-0.0476	-0.0392
	(0.0591)	(0.0504)	(0.0516)	(0.0513)
Recent Change Job	-0.0288	-0.0231	-0.0401	-0.0336
	(0.0462)	(0.0487)	(0.0475)	(0.0491)
Parent Education	-0.0511	-0.0260	-0.0109	-0.0298
	(0.0563)	(0.0516)	(0.0577)	(0.0505)
Spouse Has Job	-0.0261	-0.0223	-0.0698*	-0.0688*
	(0.0420)	(0.0399)	(0.0421)	(0.0404)
Observations	7,000	6,501	5,637	5,294
Sample	All	All	Urban Hukou	Urban Hukou

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$